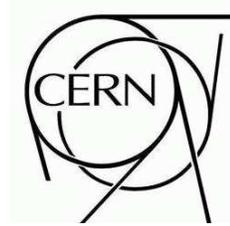# ATLAS NOTE

August 25, 2009

Prospects for measuring top pair production in the dilepton channel with early ATLAS data at $\sqrt{s} = 10$ TeV

The ATLAS Collaboration

Abstract

We present prospects for measuring the $t\bar{t}$ production cross-section with the ATLAS detector at a center-of-mass energy of 10 TeV with two leptons in the final state (ee, $\mu\mu$ and $e\mu$). The measurement aims to reestablish the top signal at the LHC with first data and gain confidence with signatures that involve leptons, jets and missing transverse energy. We introduce data-driven background estimation techniques for the dominant non-top Standard Model backgrounds and define control samples for the top dilepton sample. Assuming the Standard Model and an uncertainty of 20% in the early luminosity determination, we expect a measurement of the cross-section for an integrated luminosity of $200\,\mathrm{pb}^{-1}$ in the combined channel with a relative uncertainty of $3.1(\mathrm{stat})^{+9.6}_{-8.7}(\mathrm{syst})^{+26.2}_{-17.4}(\mathrm{lumi})\%$.

1   Introduction

The measurement of the top quark pair production cross-section ($\sigma_{t\bar{t}}$) at LHC in its different sub-channels is interesting for various reasons. Uncertainties on the theoretical predictions are now at the level of 10% and a comparison with measurements will allow to test the expectations from Standard Model (SM) QCD. The abundant sample which will be available is also well suited for use as a calibration tool of the reconstructed objects in the detector and algorithms, in particular at the beginning of data taking. Furthermore $t\bar{t}$ events are often an important background in the searches for the Higgs boson or physics beyond the SM and it is thus of paramount importance to study this process in detail. Finally, cross-section measurements are also an important test of possible new production mechanisms, as a non-SM top quark production can lead to a significant increase of the cross-section. New physics may affect the cross-sections differently in various decay channels.

In this note we study the potential of the ATLAS detector for a measurement of the $\sigma_{t\bar{t}}$ with final states with two leptons (e and μ). It is assumed that the branching ratio $BR(t \rightarrow Wb) = 1$ and leptonic τ decays are included in the signal definition. The branching ratio [1] for decay modes to ee, $\mu\mu$ and $e\mu$ final states is 6.4%. For the first run we assume a center-of-mass energy $\sqrt{s} = 10$ TeV and an integrated luminosity of 200 pb^{-1}.

Top-pair production at the LHC can occur well beyond threshold giving us an opportunity to study the top quark in large quantity as well as broad phase space. A significant fraction of the $t\bar{t}$ production therefore originates from higher order processes and the availability of calculations for higher order QCD correction is vital. The LO and NLO [2, 3, 4] inclusive $\sigma_{t\bar{t}}$ calculations [5] have been extended to higher order processes. Improvements on the NLO $\sigma_{t\bar{t}}$ calculation have been achieved by including soft-gluon corrections by re-summation of large Sudakov logarithms to NLL accuracy [6]. Recently, the prescription was extended to include terms of NNLL accuracy and based on this result, "approximate" NNLO QCD predictions have been derived [7]. The uncertainty due to the mass of the top quark is small, now that the precision of the top mass measurement at the Tevatron is about 1%. It is also noted [7], that the sensitivity of the total cross-section to soft-gluon emission is somewhat different at Tevatron and LHC as the production at Tevatron is largely dominated by parton kinematics close to threshold.

For the first LHC data, the beam energy is expected to be lower than the designed 7 TeV. For this study we assume the beam energy to be 5 TeV. This reduces the $t\bar{t}$ cross-section at $\sqrt{s} = 10$ TeV by 55% compared to $\sqrt{s} = 14$ TeV, giving a value of 400 pb with an uncertainty of approximately 11 % at NLO and 6 % at NNLO, dominated by renormalization and factorization scales and knowledge of the parton distribution functions. For an integrated luminosity of 200 pb^{-1}, we expect about 27 thousand single-lepton events and 5 thousand dilepton events.

2   Object Reconstruction and Event Selection

The general selection of $t\bar{t}$ dilepton (ee, $\mu\mu$ and $e\mu$) candidate events requires a positive decision of a single high-p_T lepton trigger (Section 2.1). The events are required to have two oppositely-charged lepton (e or μ) candidates with a high transverse momentum of $p_\mathrm{T} > 20$ GeV (Sections 2.2 and 2.4, respectively). The $t\bar{t}$ dilepton events are expected to have large missing transverse energy $E_\mathrm{T}^\mathrm{miss}$ (Section 2.5), whereas the Drell-Yan events (one of the largest background in the ee and $\mu\mu$ channels) are expected to have low $E_\mathrm{T}^\mathrm{miss}$. We also require an event to contain at least 2 jets (Section 2.3) with $p_\mathrm{T} > 20$ GeV that mainly arise from the hadronised b-quarks. The object selections were determined carefully so that the analysis will not rely on variables that are difficult to calibrate with a small amount of data.

2.1 Trigger

The low luminosity conditions that are expected to be prevalent during the early running period at a center-of-mass energy of 10 TeV are favorable for efficiently triggering on top-quark events with leptons in the final state. Single-lepton triggers with low threshold and loose selection requirements can run without any pre-scaling and can by themselves provide a high acceptance $t\bar{t}$ trigger.

During the early running period, a large fraction of the 200 pb^{-1} of data is expected to be collected with peak luminosities of around 10^{32} cm^{-2}s^{-1}. The cross-section analysis described in this paper hence assumes the collection of the $t\bar{t}$ signal events using a single-electron and a single-muon trigger with a p_T threshold of 15 GeV. The electron candidates considered for the trigger are required to exhibit a good shower shape and to have an associated inner detector track but there is no explicit isolation requirement.

The trigger efficiencies for the single-lepton triggers can be measured from the data itself using Z boson decays to leptonic final states. One method of trigger extraction from data is tag-and-probe, where one obtains an unbiased "probe" electron exploiting the Z peak [8]. To be applied to $t\bar{t}$ events, the efficiencies should be measured as a function of p_T, η, hadronic activity of the event, isolation of leptons. The total uncertainty is expected to be of the order of 1% during the early running period. The cross-section analysis in this paper uses the single-lepton trigger efficiencies as predicted from the trigger simulation (Fig. 1).
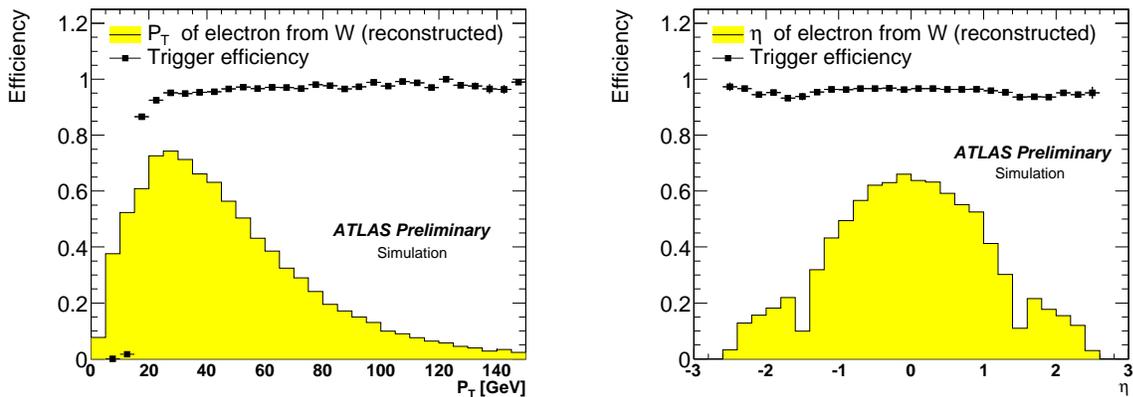

Figure 1: Electron trigger efficiency as a function of p_T (on the left) and η (on the right) for the signal events. The shaded area shows the p_T and η distributions of reconstructed electrons that match true electrons from W decays. The efficiencies are measured with respect to the offline event selection (Section 2.2).

2.2 Electrons

Electrons are identified in the electromagnetic calorimeter by matching high momentum inner detector tracks to high-energy EM clusters. A cut-based quality requirement identifies signal electrons with high efficiency while keeping the misidentification ("fake") rate from hadronic jets low; the electrons are required to satisfy the medium purity cuts [8]. We require electron $E_T > 20$ GeV; an additional requirement of isolation $E_T < 6$ GeV in the $\Delta R_{\eta-\phi}$ cone of 0.2 further reduces the fake rate. To maintain uniform efficiency across the accepted regions, we require $|\eta| < 2.47$. We remove electrons in a region $(1.37 < |\eta| < 1.52)$ where a significant amount of dead material in front of the calorimeter is

present, and the efficiency is generally lower while the fake rate is higher. Electron efficiency plots are shown in Fig. 2.
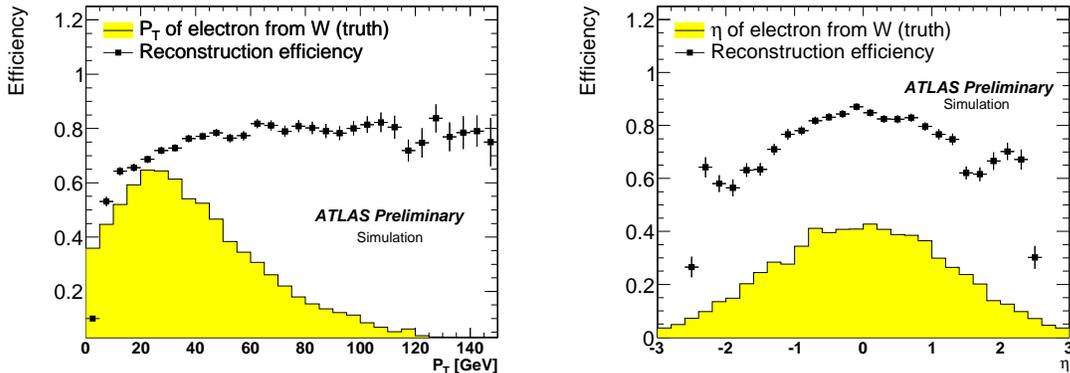

Figure 2: Electron reconstruction efficiencies as a function of electron p_T (left) and η (right) for signal events as predicted from the simulation. The shaded area shows the p_T and η distributions of true electrons from W decays.

2.3 Jets

The ATLAS "cone" algorithm with radius $\Delta R_{\eta-\phi} = 0.4$ was used to reconstruct jets based on the calorimeter tower input calibrated using "H1" cell weights [8]. Jets are required to have $|\eta| < 2.5$ and a transverse momentum of at least 20 GeV. We do not require jets to be b-tagged since it has been shown that b-tagging is not required for initial measurements of the $t\bar{t}$ cross-section [8], and it will take some time to understand and calibrate with the initial data. Since identified electrons are also reconstructed as jets, those jets near selected electrons are removed to avoid double-counting. For this we remove a jet if there is an identified electron within $\Delta R < 0.2$.

2.4 Muons

We require muons to be reconstructed by both the muon chambers and the inner detector in the region of $|\eta| < 2.5$. A transverse momentum cut at 20 GeV is applied to avoid the trigger turn-on region and the isolation E_{T} is required to be smaller than 6 GeV in a $\Delta R_{\eta-\phi}$ cone of 0.2. The main source of fake muons are the muons originating from heavy-flavor decay, so we require that there is no identified jet within $\Delta R_{\eta-\phi} < 0.3$ of a muon. Muon efficiency plots are shown in Fig. 3.

2.5 Missing Transverse Energy

Missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) is associated with particles that escape detection. For example, $E_{\mathrm{T}}^{\mathrm{miss}}$ is the signature of weakly interacting neutral particles such as neutrinos. It can also come from mismeasurement of the true E_T of the objects, or from backgrounds such as cosmic rays or beam halo. Missing E_{T} ($\vec{\not{E}}_T$) is defined by $\vec{\not{E}}_T = -\sum_i E_T^i \hat{n}_i$, where i is the calorimeter tower number for $|\eta| < 4.9$, and $\hat{n}_i$ is a unit vector perpendicular to the beam axis and pointing at the ith tower. We define the magnitude $E_{\mathrm{T}}^{\mathrm{miss}} = |\vec{\not{E}}_T|$. Corrections are made to the $E_{\mathrm{T}}^{\mathrm{miss}}$ for the reconstructed objects, including electrons, muons and jets. In addition, the $E_{\mathrm{T}}^{\mathrm{miss}}$ is also corrected for dead material in the cryostat.
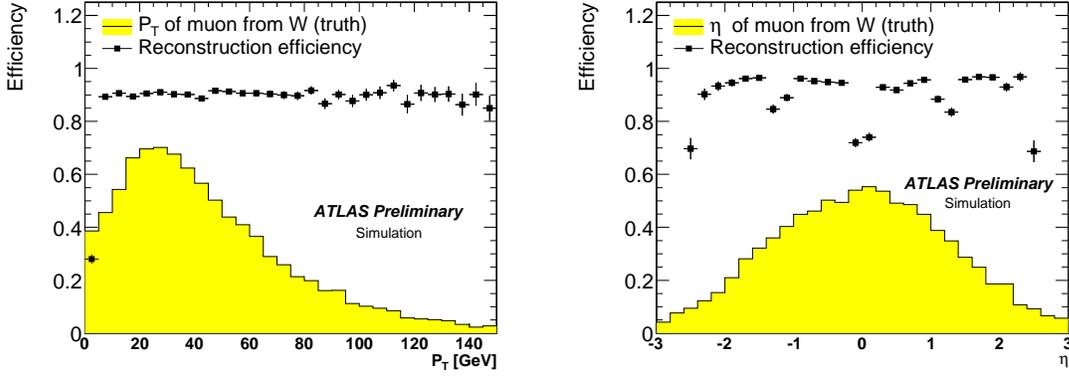
Figure 3: Muon reconstruction efficiencies as a function of muon p_T (left) and η (right) for signal events as predicted from the simulation. The shaded area shows the p_T and η distributions of true muons from W decays.

The optimization at $\sqrt{(s)} = 14$ TeV [8] for the best significance $S/\sqrt{S+B}$ determined a cut of $E_{\mathrm{T}}^{\mathrm{miss}} > 35$ GeV which rejects a large fraction of $Z \to ee$ $(\mu\mu)$ events for the ee $(\mu\mu)$ channel. To further suppress Drell-Yan background for the ee and $\mu\mu$ channels, the invariant dilepton mass is required to be inconsistent with the Z mass, $|m_Z - m_{\ell^+\ell^-}| > 5$ GeV. The $e\mu$ channel is expected to yield the best signal-to-background ratio, since it does not suffer from Drell-Yan background. The $E_{\mathrm{T}}^{\mathrm{miss}}$ cut was optimized [8] for the best significance and found to be $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV.

3 Background and Signal Expectation

Several processes can mimic the $t\bar{t}$ dilepton signal by producing the same signature in the detector (two leptons, $E_{\mathrm{T}}^{\mathrm{miss}}$ and jets), but with one or more misidentified objects. One example is $Z \to ee +$ jets or $Z \to \mu\mu +$ jets events, which can have significant $E_{\mathrm{T}}^{\mathrm{miss}}$ because of resolution effects or mismeasured jet energies (Section 3.1.1). Another example is jets misidentified as leptons in QCD and W+jets events, which can mimic the $t\bar{t}$ dilepton signature (Section 3.1.2).

Additionally, a few processes can produce two real leptons, a neutrino, and a number of jets and thus get mistaken for the $t\bar{t}$ dilepton signal. Events with two gauge bosons (WW, WZ and ZZ) or single-top (Wt-channel) production are an example. These backgrounds are considered small and are calculated based on MC efficiencies (Section 3.2). We also describe the $t\bar{t}$ acceptance calculation in Section 3.2.

3.1 Data-Driven Methods for Background Determination

We take a data-driven approach to some backgrounds due to difficulties in the MC simulation associated with modeling them, for instance the QCD and the Drell-Yan (DY) contributions. We also cannot trust the MC simulation alone to predict instrumental effects and backgrounds involving the tails of distributions. In this Section we describe procedures to estimate background to dileptons from the DY events and from jets misidentified as leptons.

3.1.1 Drell-Yan Background

Requiring large E_T^{miss} and a dilepton mass outside the Z window significantly cuts away the DY background, but does not completely eliminate it. The DY events in the tail of the E_T^{miss} distribution and away from the Z peak still pass our event selection. We use a data-driven technique to estimate the number of these events; the DY contribution to our signal region is estimated by scaling the MC prediction to match the observed number of events in sideband regions of the data. The method is applied separately for ee and $\mu\mu$ events.
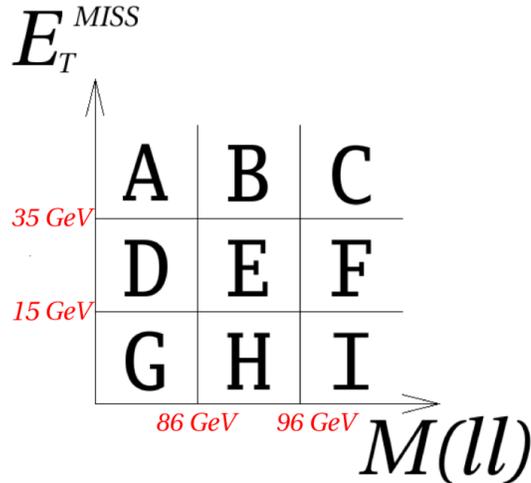

Figure 4: Diagram of E_T^{miss} versus $M_{\ell^+\ell^-}$ regions used for Drell-Yan data-driven background estimates.

Using the grid shown in Fig. 4, we define the columns by the Z mass window between 86 GeV and 96 GeV. We select dilepton ee and $\mu\mu$ events using the selection criteria for leptons described in Sections 2.2 and 2.4. The rows are defined by E_T^{miss} values of 15 GeV and 35 GeV. We populate the nine bins using a sample of pseudo-data (which contains $t\bar{t}$, single top, dibosons and W/Z+jets) and ALPGEN Z MC samples. The relative admixture of physical processes are different in the different bins. For instance, the $Z \to \ell\ell$ signal falls largely in bin "H," whereas $t\bar{t}$ dilepton events contribute to bins "A", "B" and "C". Using the formulas in Eq. 1, we estimate the amount of DY background in our signal region ("A" and "C") by scaling the MC DY distribution to match the pseudo-data in both the low-E_T^{miss} region as well as in the Z peak:

$$A_{Est} = G_{Data}(\frac{A_{MC}}{G_{MC}})(\frac{B_{Data}}{H_{Data}})(\frac{H_{MC}}{B_{MC}}), \qquad C_{Est} = I_{Data}(\frac{C_{MC}}{I_{MC}})(\frac{B_{Data}}{H_{Data}})(\frac{H_{MC}}{B_{MC}}) \qquad (1)$$

In the above equations, the subscripts "Data" and "MC" correspond to the pseudo-data and inclusive Z MC samples, respectively. The total DY background estimate is given by $A_{Est} + C_{Est}$. We obtain a final measurement as well as a systematic uncertainty on the estimate by varying both the E_T^{miss} boundaries and the Z window (see Section 5.8).

3.1.2 Jets Misidentified as Leptons

W+jets and QCD events with jets misidentified as leptons can mimic the $t\bar{t}$ dilepton signature, and therefore contribute a significant background to our selection. Due to difficulties in accurately simulating events with jets misidentified as leptons, we estimate this contribution to our data sample in a data-driven

way. There are two types of dilepton events with fake leptons: events with both a real and a fake lepton (W+jets) and events with two fake leptons (QCD). Events with two fake leptons typically have small E_T^{miss}, so for this analysis they are neglected. The probability for a jet to fake a lepton can be determined by defining an auxiliary loose lepton selection, and measuring the efficiencies of both this loose and the default (tight) lepton selection in two independent event samples, dominated by real and fake leptons respectively.

We define a selection of loose leptons, such that most loose lepton candidates are actually misidentified jets and only a small fraction are signal leptons. We define probabilities for a real and a fake lepton, ϵ and f, respectively, to be reconstructed as a tight lepton

$$\epsilon = \frac{N_{\text{tight,real}}}{N_{\text{tight,real}} + N_{\text{loose,real}}}, f = \frac{N_{\text{tight,fake}}}{N_{\text{tight,fake}} + N_{\text{loose,fake}}} \tag{2}$$

We measure ϵ using the tag-and-probe method in events with two leptons with $M(\ell, \ell)$ within 5 GeV of the Z boson mass and $E_T^{miss} < 15$ GeV. We measure f separately in two samples dominated by fakes. The first one is a "low $\Delta\phi$" sample requiring one tight or loose lepton, $E_T^{miss} > 15$ GeV, the angle between the lepton and the E_T^{miss}, $\Delta\phi < 1$ rad. The second one is a "low-E_T^{miss}" sample requiring one tight or loose lepton and $E_T^{miss} < 15$ GeV. These two selections are orthogonal and are useful for estimating our systematic uncertainty associated with applying the fake rate measured in the single lepton sample to the dilepton sample.

The "low $\Delta\phi$" selection rejects 75% (80%) of W's in an inclusive W MC sample for the muons (electrons). The "low E_T^{miss}" selection rejects 98% of W's in an inclusive W MC sample. A signal subtraction is employed to remove the remaining signal events from these samples. We use the "low E_T^{miss}" sample for our fake predictions, the results are shown in Figure 5. Note that while the final estimates for the predicted number of fake events passing the dilepton selection criteria should not be a sensitive function of the choice of loose lepton definition, the tight fake probabilities are sensitive to this choice.
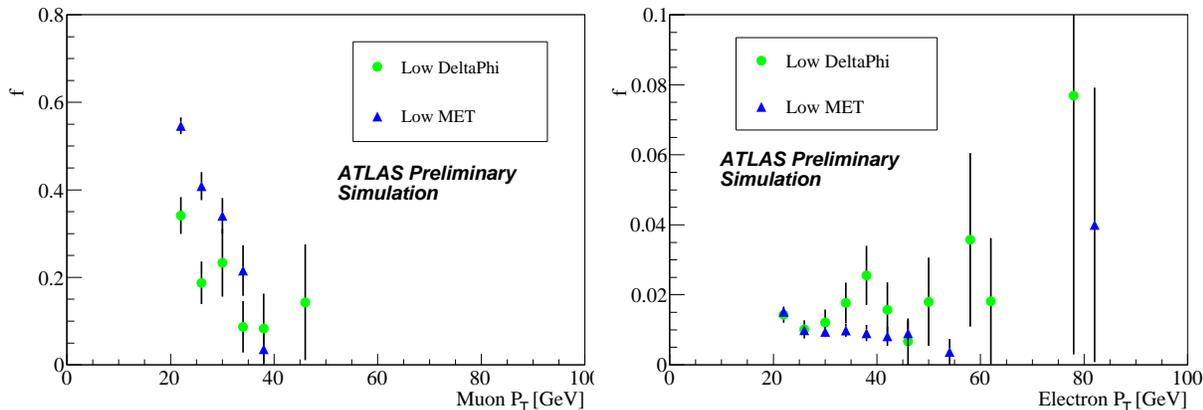
Figure 5: Tight fake probabilities f measured as a function of p_T for muons (left) and electrons (right) in a multijet sample.

We extract the number of fake events in our selection using the efficiencies and fake probabilities. We consider three different types of events at the reconstruction level: events with two tight leptons, TT, events where the highest p_T lepton is tight (loose) and second lepton is loose (tight) TL (LT). At the truth level there are three different sources of events: events with two real leptons, RR, events where the highest p_T lepton is real and the second lepton comes from a jet, RF, and events where the highest p_T

lepton comes from a jet and the second lepton is real, FR. The two sets of events are related to each other with an efficiency matrix:

$$\begin{bmatrix} N_{\text{TT}} \\ N_{\text{TL}} \\ N_{\text{LT}} \end{bmatrix} = \begin{bmatrix} \epsilon_1\epsilon_2 & \epsilon_1 f_2 & f_1\epsilon_2 \\ \epsilon_1(1-\epsilon_2) & \epsilon_1(1-f_2) & f_1(1-\epsilon_2) \\ (1-\epsilon_1)\epsilon_2 & (1-\epsilon_1)f_2 & (1-f_1)\epsilon_2 \end{bmatrix} \begin{bmatrix} N_{\text{RR}} \\ N_{\text{RF}} \\ N_{\text{FR}} \end{bmatrix} \qquad (3)$$

where ϵ_1 and ϵ_2 (f_1 and f_2) are the efficiencies for real (fake) first and second leptons to be reconstructed as a tight lepton.

This equation when inverted yields the number of events coming from each source of events. All events that are reconstructed with two tight leptons and do not come from two real leptons are considered to be fakes. This yields the following expression:

$$N_{\text{Fake}} = \left[\frac{f_2(\epsilon_2-1)}{\epsilon_2-f_2} + \frac{f_1(\epsilon_1-1)}{\epsilon_1-f_1} \right] N_{\text{TT}} + \frac{f_2\epsilon_2}{\epsilon_2-f_2} N_{\text{TL}} + \frac{f_1\epsilon_1}{\epsilon_1-f_1} N_{\text{LT}} \qquad (4)$$

As expressed above, this technique assumes a constant efficiency and tight fake probability. However, it is likely that these will be dependent on various kinematic variables (such as p_{T} and η). The technique is easily extended if the efficiencies (ϵ_1, ϵ_2) and fake probabilities (f_1, f_2) are measured as functions of kinematic variables (such as p_{T} and η). In this case, rather than deriving the fake estimate by weighting event counts ($N_{\text{TT}}, N_{\text{TL}}, N_{\text{LT}}$), the fake estimate is expressed as a sum of weighted events, using the coefficients of ($N_{\text{TT}}, N_{\text{TL}}, N_{\text{LT}}$) in Eq. 4 as the weight for TT, TL or LT events, respectively. This approach has the advantage of providing a coherent prediction in all kinematic variables, rather than a prediction for a specific histogram. For example, one can describe the fake estimate as a function of the number of reconstructed jets in each channel, see Fig. 6.

The technique outlined above should be sufficient for determining the jets faking lepton background to our selection. Only once we are able to look at the control samples in real data it will become apparent what the fake rates are in reality, and what the systematics on these fake rates are (Section 5.9). Once these are available the outlined technique can be used to predict this background.

3.2   Monte Carlo Based Determination of Signal and Background

The expected contributions from signal and from remaining background processes (single top, diboson, $Z \to \tau\tau$, $Wb\bar{b}$) are calculated using MC methods. We also estimate contributions from W+jets and DY MC samples, for which systematic uncertainties are obtained from the data-driven methods described in Section 3.1. The MC sets used in this analysis have passed through a full simulation of the ATLAS detector. Tables 1-3 summarize the estimated contributions for the ee, $\mu\mu$ and $e\mu$ channels.

Many of these processes have two high-p_{T}, prompt leptons which are reconstructed. Several of them, however, may produce misidentified leptons, from light- or heavy-flavor jets. For example, production of Z events in association with jets can produce such an event if one of the Z decay products is lost and a jet is misidentified as a lepton. This portion of the contribution is already described by the data-driven estimate (Section 3.1).

To avoid double counting, we require the $t\bar{t}$ dilepton signal and the MC-driven portion of the background to have two real prompt leptons in the truth-record, matched to the reconstructed leptons. In particular, for the $t\bar{t}$ dilepton acceptance calculation the two leptons are required to be generated from a $t \to W \to \ell$ or a $t \to W \to \tau \to \ell$ decay chain. Figure 7 shows that the angular distance between reconstructed and true leptons is very small; we chose a matching distance of $\Delta R < 0.05$.

With this procedure the background contribution in Tables 1-3 have to be increased by those events in the $t\bar{t}$ dilepton sample that did not pass this truth-matching cut. On average $\sim 2\%$ of the selected $t\bar{t}$ dilepton events were found to be in this category.
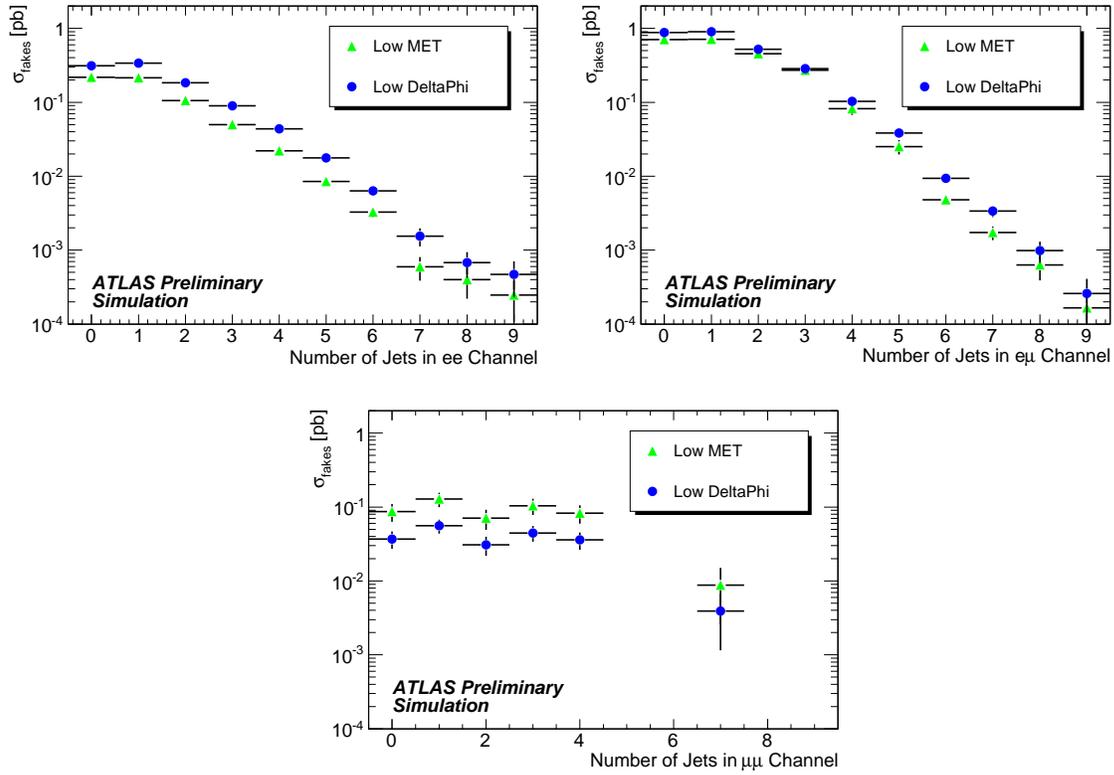
8

Figure 6: Fake predictions for ee (top left), $e\mu$ (top right), and $\mu\mu$ (bottom) as a function of the number of jets.
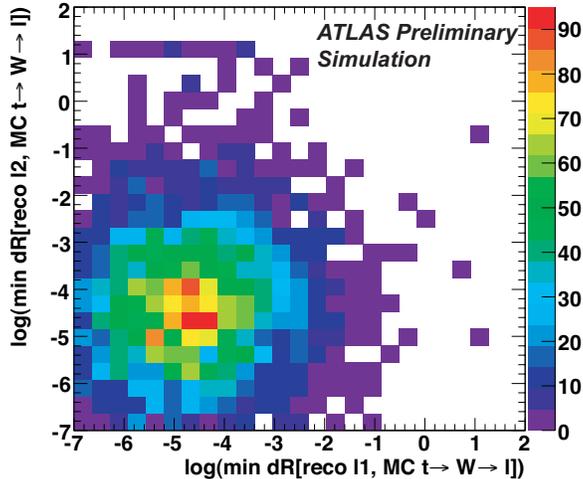
Figure 7: Minimum angular distance in ΔR between reconstructed leptons and leptons from the truth record which originate from $t \to W \to l$ decays. Overflow is shown at $\log(\Delta R) = 1$.

The backgrounds are estimated using theoretical cross-sections. The calculated number of expected background events from a process is given by

$$N_{expected} = \mathcal{L} \times \sigma_{process} \times \frac{N_{passed}}{N_{total}} \qquad (5)$$

where acceptance $A = \frac{N_{passed}}{N_{total}}$ is a ratio of events which pass selection criteria. We take into account weights w_i of the MC events, $N_{passed} = \sum_{i \in \{passed\ events\}} w_i$, $N_{total} = \sum_{i \in \{all\ events\}} w_i$. For real data we will also correct for efficiency differences between MC and data. This is implemented by weighting MC events with appropriate object identification and trigger efficiencies scale factors. The scale factors are defined as $\epsilon^i_{data}/\epsilon^i_{MC}$, where ϵ^i_{data} are identification (ID) and trigger efficiencies measured in data from Z events, and ϵ^i_{MC} are efficiencies measured in inclusive Z MC. The acceptance is a convolution of kinematic, geometric and reconstruction efficiency contributions.

Dilepton branching ratios include electrons, muons and leptonic tau decays of the W bosons in the top pair decay. We estimate BR($t\bar{t} \to ee$) = (1.67 ± 0.05)%, BR($t\bar{t} \to \mu\mu$) = (1.64 ± 0.05)% and BR($t\bar{t} \to e\mu$) = (3.40 ± 0.10)%. We also estimate top dilepton acceptances to be A(ee) = (16.5 ± 0.4)%, A($\mu\mu$) = (26.1 ± 0.4)% and A($e\mu$) = (26.5 ± 0.3)%.

Based on MC simulation for signal and background, we present the expected kinematic distributions for an integrated luminosity of 200 pb^{-1}. Figures 8-10 show the purely MC based signal and background E_T^{miss} and jet multiplicity distributions for each sub-channel. The corresponding Tables 1-3 contain the full list of event types for each MC sample. The uncertainties are due to the limited number of events in the MC samples. Statistical uncertainties for different luminosities are shown in Table 4. These statistical uncertainties assume a Poissonian error on the total number of observed events whereas the error on the background is taken from the MC estimation. The table shows the combined statistical uncertainties.

Figure 8: Missing transverse energy distribution (left) after requiring two opposite signed leptons and jet multiplicity distribution (right) after all cuts (except the $N_{jets} > 1$ cut) for ee dilepton signal and MC based background estimations after all cuts. The samples are normalized to $200\,\mathrm{pb}^{-1}$.

Figure 9: Missing transverse energy distribution (left) after requiring two opposite signed leptons and jet multiplicity distribution (right) after all cuts (except the $N_{jets} > 1$ cut) for $\mu\mu$ dilepton signal and MC based background estimations after all cuts. The samples are normalized to $200\,\mathrm{pb}^{-1}$.
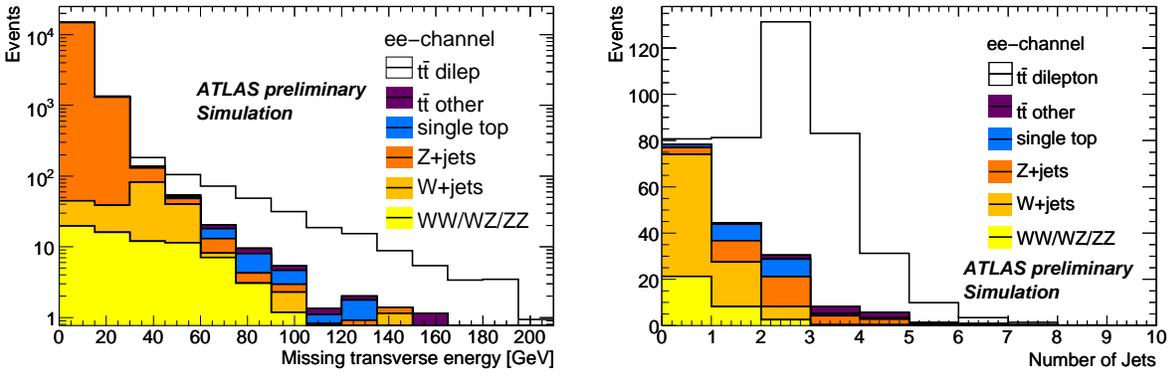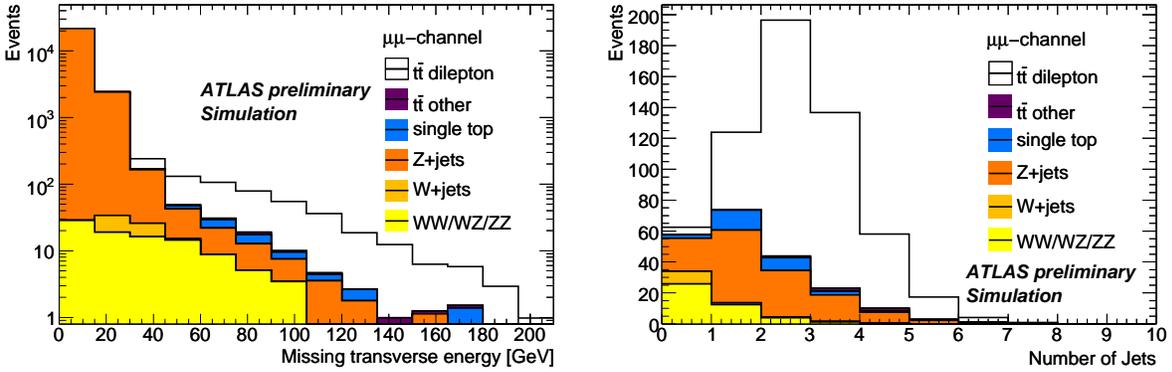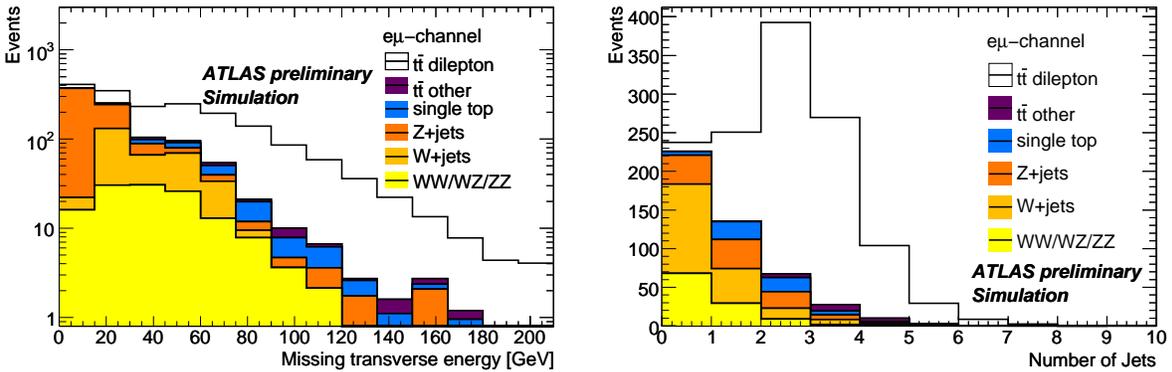

Figure 10: Missing transverse energy distribution (left) after requiring two opposite signed leptons and jet multiplicity distribution (right) after all cuts (except the $N_{jets} > 1$ cut) for $e\mu$ dilepton signal and MC based background estimations after all cuts. The samples are normalized to $200\,\mathrm{pb}^{-1}$.

Table 1: Expected number of selected events for the ee channel selection for $200\,\mathrm{pb}^{-1}$. The errors are only shown for the last two columns for readability, but they are fully taken into account in the summation and the S/B calculation. The column labeled "true" represents the effect of applying truth-matching cuts. The column labeled "fake" shows the number of events which failed the truth-matching cuts.

	lepton sel.	inv. mass cut	$E_\mathrm{T}^\mathrm{miss}$cut	jet cut	trigger	true	fake
$t\bar{t}$ dilepton	351	322	261	220	214 ± 6	209 ± 6	4
$t\bar{t}$ other	15	13	10	9	8 ± 1	$0^{+0.1}_{-0}$	8
single top	27	25	18	9	9 ± 2	7 ± 1	2
$Z \to ee$	68231	16283	24	14	13^{+2}_{-1}	11^{+2}_{-1}	2
$Z \to \tau\tau$	156	154	10	7	7^{+2}_{-1}	6^{+2}_{-1}	1
$W \to e\nu$	126	118	56	7	7^{+4}_{-2}	0^{+4}_{-0}	7
$W \to \tau\nu$	7	7	7	1	1^{+4}_{-1}	0^{+4}_{-0}	1
diboson	145	73	33	3	3 ± 1	2 ± 1	1
sum bkg	68707	16673	157	51	49^{+8}_{-3}	$54^{+11}_{-3}{}^{*}$	
S/B	0.0	0.0	1.7	4.3	$4.3^{+0.7}_{-0.3}$	$3.9^{+0.8}_{-0.3}$	
$S/\sqrt{S+B}$	1.3	2.5	12.5	13.4	13.2	12.9	

* Including all fakes.

Table 2: Expected number of selected events for the $\mu\mu$ channel selection for $200\,\mathrm{pb}^{-1}$. The errors are only shown for the last two columns for readability, but they are fully taken into account in the summation and the S/B calculation. The column labeled "true" represents the effect of applying truth-matching cuts. The column labeled "fake" shows the number of events which failed the truth-matching cuts.

	lepton sel.	inv. mass cut	$E_\mathrm{T}^\mathrm{miss}$cut	jet cut	trigger	truth	fake
$t\bar{t}$ dilepton	530	490	400	343	332 ± 8	327 ± 7	6
$t\bar{t}$ other	9	8	6	6	5 ± 1	$0^{+0.1}_{-0}$	5
single top	38	35	27	12	12 ± 2	11 ± 2	1
$Z \to \mu\mu$	109331	24716	113	49	47^{+3}_{-2}	44^{+3}_{-2}	4
$Z \to \tau\tau$	263	262	19	11	10^{+2}_{-1}	10^{+2}_{-1}	0
$W \to \mu\nu$	14	14	7	0	0^{+4}_{-0}	0^{+4}_{-0}	
$W \to \tau\nu$	1	1	0	0	0^{+4}_{-0}	0^{+4}_{-0}	
Wbb	1	1	1	0	$0.4^{+0.2}_{-0.1}$	$0^{+0.2}_{-0}$	
diboson	211	100	45	5	5 ± 1	5 ± 1	1
sum bkg	109868	25137	218	84	81^{+8}_{-3}	$87^{+11}_{-3}{}^{*}$	
S/B	0.0	0.0	1.8	4.1	$4.1^{+0.4}_{-0.2}$	$3.8^{+0.5}_{-0.2}$	
$S/\sqrt{S+B}$	1.6	3.1	16.1	16.6	16.4	16.1	

* Including all fakes.

4 Analysis Strategy

4.1 Cross-Section Determination

The expected integrated luminosity in the first year is $200\,\mathrm{pb}^{-1}$ and it will be known with a relatively large ($\sim 20\%$) uncertainty. It is important that in estimating the expected uncertainty on the cross-section,

Table 3: Expected number of selected events for the $e\mu$ channel selection for $200\,\mathrm{pb}^{-1}$. The errors are only shown for the last two columns for readability, but they are fully taken into account in the summation and the S/B calculation. The column labeled "true" represents the effect of applying truth-matching cuts. The column labeled "fake" shows the number of events which failed the truth-matching cuts.

	lepton sel.	$E_\mathrm{T}^\mathrm{miss}$cut	jet cut	trigger	truth	fake
$t\bar{t}$ dilepton	908	845	715	698 ± 11	683 ± 11	15
$t\bar{t}$ other	24	21	21	20 ± 2	$0^{+0.1}_{-0}$	20
single top	59	55	26	25 ± 3	23 ± 3	2
$Z \to \mu\mu$	110	15	4	4^{+1}_{-0}	0^{+1}_{-0}	4
$Z \to \tau\tau$	421	97	28	27^{+3}_{-2}	25^{+3}_{-2}	2
$W \to e\nu$	14	11	2	2^{+4}_{-1}	0^{+4}_{-0}	2
$W \to \mu\nu$	165	144	20	17^{+6}_{-4}	0^{+4}_{-0}	17
$W \to \tau\nu$	11	9	2	2^{+4}_{-1}	0^{+4}_{-0}	2
Wbb	1	1	0	$0.4^{+0.3}_{-0.1}$	$0^{+0.2}_{-0}$	
diboson	135	111	11	10 ± 1	10 ± 1	0.2
sum bkg	939	464	113	108^{+10}_{-6}	123^{+12*}_{-6}	
S/B	1.0	1.8	6.3	$6.5^{+0.6}_{-0.3}$	$5.6^{+0.6}_{-0.3}$	
$S/\sqrt{S+B}$	21.1	23.4	24.9	24.6	24.1	

* Including all fakes.

Table 4: Statistical error and significance on the cross-section σ for the different sub-channel from MC based signal and backgrounds. The error on the number of observed events is Poissonian and the error on the background estimation is taken from the statistical error in the MC sample.

statistical error $\Delta\sigma/\sigma$ [%]	ee	$\mu\mu$	$e\mu$
$\mathcal{L} = 50\,\mathrm{pb}^{-1}$	21.7	15.7	9.9
$100\,\mathrm{pb}^{-1}$	13.2	9.9	6.4
$200\,\mathrm{pb}^{-1}$	8.5	6.6	4.3
significance $S/\sqrt{S+B}$	ee	$\mu\mu$	$e\mu$
$\mathcal{L} = 50\,\mathrm{pb}^{-1}$	6.6	8.2	12.3
$100\,\mathrm{pb}^{-1}$	9.3	11.6	17.4
$200\,\mathrm{pb}^{-1}$	13.2	16.4	24.6

our approach incorporates various systematic uncertainties expected in the first year.

The measurement is based on a simple counting experiment, thus we model the observed count N^{obs} as being Poisson distributed about some expectation N_{tot}^{exp}. The tot subscript indicates that there are several contributions: i. e. the signal and various backgrounds (indexed by k).

$$Pois(N^{obs}|N_{tot}^{exp}) = Pois(N^{obs}| \sum_{k\in\{sig,bkg\}} N_k^{exp}) \tag{6}$$

For each of these contributions, the expectation is simply a product of the integrated luminosity, $\mathcal{L}$, the

cross-section for the process, σ_k, and several efficiencies (indexed by j), ε_{jk}[1]. Thus we have:

$$N_k^{exp} = \mathcal{L}\sigma_k \prod_j \varepsilon_{jk} \,. \tag{7}$$

Clearly, the most likely value of σ_{sig} will be the one which realizes $N^{obs} = N_{tot}^{exp}$. This leads to the familiar result:

$$\hat{\sigma}_{sig} = \frac{N^{obs} - \sum_{k\in\{bkg\}} N_k^{exp}}{\mathcal{L}\prod_j \varepsilon_{j\,sig}} \,, \tag{8}$$

where $\hat{\sigma}$ denotes the maximum likelihood estimate of σ. The intuitive form of Eq. 8 motivates techniques based on background-subtraction; however, that approach makes it less clear how to estimate the uncertainty on the measured value of σ. These complications become more severe when systematics are included (for example, the large uncertainty on $\mathcal{L}$, which appears in the denominator, will no longer be symmetric and it is not uncorrelated to σ). For these reasons, it is preferred to model the observation directly (i. e. Eqs 6 and 7) without background subtraction. Regardless, Eq. 8 remains valid with or without background subtraction.

To incorporate uncertainty on the luminosity, the likelihood function is extended to

$$L(\sigma_{sig}, \mathcal{L}) = Pois(N^{obs}|N_{tot}^{exp}) \times Gaus(\tilde{\mathcal{L}}|\mathcal{L}, \Delta_{\mathcal{L}}) \,, \tag{9}$$

where $\mathcal{L}$ is interpreted as the true, unknown integrated luminosity, $\tilde{\mathcal{L}}$, is the nominal estimate of the luminosity, and $\Delta_{\mathcal{L}}$ is the uncertainty on that estimate.

Similarly, the likelihood function can be extended to incorporate uncertainties on the efficiencies ε_{jk}. This is achieved in two steps. First, we group the sources of systematics, α, such that the corresponding variations in ε are expected to be uncorrelated. Next, we vary the sources of the systematics (e. g. jet energy scale, trigger efficiencies, etc.) by the $\pm 1\sigma$ variations and determine the $\varepsilon_{jk}(\alpha_j^{\pm})$ for each signal and background. A change in the source of the jth systematic will cause a totally correlated variation among the contributions (indexed by k). Thus we describe the efficiency as a piece-wise linear function $\varepsilon_{jk}(\alpha_j)$, and this parametrized efficiency is used to in place of the nominal efficiency. Additional Gaussian terms are added to the likelihood function to represent the constraints on the α_j derived from auxiliary measurements or our assumptions about the uncertainty in the Monte Carlo modeling. This modeling of the likelihood function is repeated for each of the channels, and the total likelihood function is simply the product of the individual likelihood functions. Several of the α_j are shared between the channels (for example, the jet energy scale uncertainty), which explicitly introduces correlations between the channels. Thus, we arrive at the final likelihood function:

$$L(\sigma_{sig}, \mathcal{L}, \alpha_j) = \prod_{l\in\{ee,\mu\mu,e\mu\}} \left\{ \prod_{i\in bins} \left[Pois(N_i^{obs}|N_{i,tot}^{exp})\, Gaus(\tilde{\mathcal{L}}|\mathcal{L}, \sigma_{\mathcal{L}}) \prod_{j\in syst} Gaus(0|\alpha_j, 1) \right] \right\} \,. \tag{10}$$

4.2 Extracting Measurements from the Profile Likelihood Ratio

The likelihood function can be maximized to determine the maximum likelihood estimate of all the parameters $\hat{\sigma}_{sig}, \hat{\mathcal{L}}, \hat{\alpha}_j$. We then consider the likelihood ratio

$$r(\sigma_{sig}) = \frac{L(\sigma_{sig}, \hat{\mathcal{L}}, \hat{\alpha}_j)}{L(\hat{\sigma}_{sig}, \hat{\mathcal{L}}, \hat{\alpha}_j)} \tag{11}$$

---

[1] Here we include geometrical acceptances among the ε_{jk}.

and the profile likelihood ratio:

$$\lambda(\sigma_{sig}) = \frac{L(\sigma_{sig}, \hat{\hat{\mathcal{L}}}, \hat{\hat{\alpha}}_j)}{L(\hat{\sigma}_{sig}, \hat{\mathcal{L}}, \hat{\alpha}_j)} \tag{12}$$

where $\hat{\hat{\mathcal{L}}}$ and $\hat{\hat{\alpha}}_j$ represent the conditional maximum likelihood estimates of $\mathcal{L}$ and α_j holding σ_{sig} fixed. Note, the profile likelihood is always greater than the likelihood ratio, except at the maximum likelihood estimate where they are equal. This means that the curve of $-2\log\lambda$ is broader than $-2\log r$, and the difference in the intervals can be attributed to systematics.

Wilks' theorem states that under certain conditions, which are satisfied in this case, the distribution of $-2\log\lambda(\sigma_{sig}^{true})$ is asymptotically distributed as a χ^2 distribution with one degree of freedom. We utilize this relationship to establish 68%, 90%, and 95% confidence intervals. In order to estimate the expected uncertainty on the signal cross-section measurement we used a hypothetical dataset in which the number of events after all cuts was exactly the Standard Model prediction. The estimates of parameters and their uncertainties derived from hypothetical dataset have been shown to be good estimates of the mean of their respective distributions created from repeated pseudo-experiments.

5   Systematic Uncertainties

Systematic uncertainties are calculated by varying a given parameter within its uncertainty and redoing the cross-section measurement. Each systematic is described below. For illustrative purposes, in this Section we count the signal and background events to estimate $\Delta\sigma/\sigma$ separately for every systematic parameter (Eq. 13).

$$\sigma_{t\bar{t}} = \frac{N_{obs} - N_{bgd}}{A \times \epsilon \times \mathcal{L}} \tag{13}$$

where background estimates N_{bgd} are MC based N_{bkg}^{MC} (Section 3.2) and data-driven N_{bkg}^{DD} (Section 3.1): $N_{bgd} = N_{bkg}^{MC} + N_{bkg}^{DD}$.

However, to extract the uncertainty on the measured $\sigma_{t\bar{t}}$, we are using the likelihood method described in Section 4, which takes into account all correlations between different systematic uncertainties.

5.1   Luminosity ($\mathcal{L}$)

We conservatively expect a 20% luminosity systematic uncertainty for the early data. The systematic uncertainty affects the MC based part of background estimates N_{bkg}^{MC} and luminosity in the denominator of Eq. 13.

5.2   Trigger/ID Efficiencies ($\alpha_{\epsilon(e/\mu)}$)

The efficiencies will be extracted from tag-and-probe method from data, and their systematic uncertainties will be determined from varying $M_{\ell\ell}$ requirement used to define $Z \to \ell\ell$ sample. We quote 1% uncertainty on both trigger and the offline lepton identification. We conservatively consider these two efficiencies to be fully correlated, although the exact amount of correlation is yet to be estimated.

5.3   Lepton and Jet Energy Scale ($\alpha_{e/\mu ES}$ and α_{JEWS})

We estimate the uncertainty on the lepton energy scale (LES) to be $\pm$ 1%. The uncertainty due to the limited knowledge of the jet energy scale (JES) is determined by varying the energy of reconstructed jets by $\pm 5\%$ for all signal and background samples. As a first realistic scenario, the JES variation is doubled

for jets with $|\eta| > 3.2$. We correct the $E_{\mathrm{T}}^{\mathrm{miss}}$ for the LES and JES variations. The effect of LES and JES on the cross-section measurement is calculated by using the prediction of the background numbers and the selection efficiency from the LES and JES varied samples, but using the total number of observed events from the nominal sample.

5.4 PDF Uncertainties (α_{PDF})

The uncertainties due to the parton distribution function (PDF) are examined via a reweighting scheme which uses Monte-Carlo truth information about the partons that participate in the hard process [8]. The basic procedure is to evaluate the probability for a certain event with the hard partons of flavors f_1 and f_2 that have momentum fractions x_1 and x_2 and a momentum transfer Q. The probability is calculated with the PDF that was used for generation P_0 and again calculated with the new PDF P_N and the weight w for a particular event is:

$$w = \frac{P_N(f_1, x_1, Q) \times P_N(f_2, x_2, Q)}{P_0(f_1, x_1, Q) \times P_0(f_2, x_2, Q)}$$

The PDFs used for this systematic study are CTEQ6mE (20 parameters), which was used for generation, MRST2001E (15 parameters), CTEQ6.6 (22 parameters) and MRST2006nnlo (15 parameters). Only the resulting variation of the signal acceptance was propagated to the cross-section measurement. The effect on the background was neglected due to the large S/B ratio. The results are summarized in Table 5.

The final uncertainty is the largest of the uncertainties from the errors set or the difference between two central value PDFs from different PDF collaborations (here CTEQ and MRST). Since CTEQ6.6 and MRST2006nnlo are updates of the PDFs used for generation, only these uncertainties from these two PDFs are taken into account. Therefore the PDF uncertainty is largest with CTEQ6.6.

Table 5: PDF systematics on the selection efficiency ϵ for the different sub-channels. For comparison the Monte-Carlo statistical error on the efficiency is also given.

$\Delta\epsilon/\epsilon$ (=$\Delta\sigma/\sigma$) [%]	ee	$\mu\mu$	$e\mu$
statistical $\Delta\epsilon$	2.0	1.6	1.1
CTEQ6mE	3.0	2.0	2.4
MRST2001E	1.1	0.7	0.8
CTEQ6.6	2.2	1.5	1.7
MRST2006nnlo	0.6	0.3	0.4
difference between PDFs			
MRST2001E to CTEQ6mE	2.0	1.3	1.5
CTEQ6.6 to CTEQ6mE	-0.3	-0.3	-0.2
MRST2006nnlo to CTEQ6.6	-0.5	-0.3	-0.5

5.5 Initial and Final State Radiation (α_{IFSR})

Initial (ISR) and final state (FSR) radiation can affect the jet multiplicity of both signal and background events. To study this, we can either compare different models or vary the parameters of one model. The latter strategy has been studied and Λ_{QCD} and p_{T} cut off for ISR and FSR have been identified as relevant parameters. The parameters were varied in correlation in such a way to maximize and minimize the effect on the measured top mass [8]. The effect of ISR/FSR variation was estimated by calculating signal acceptance with samples generated with varied parameters. The effect on the background was

neglected due to the large S/B ratio. The systematic uncertainties are estimated to be 5.2% for the ee, 4.0% for the $\mu\mu$ and 2.2% for the $e\mu$ channel.

5.6   Monte-Carlo Model ($\alpha_{MC\ model}$)

The uncertainty on kinematic distributions predicted by Monte-Carlo simulation can alter both signal and background acceptance. To estimate this uncertainty, we compare predictions from different generators, MC@NLO, AcerMC (LO), and Alpgen. We estimate the uncertainty on the acceptance due to different Monte-Carlo models to be 10% for all the channels.

5.7   Theoretical Cross-Section ($\alpha_{cross\ sec.}$)

While data-driven techniques are preferred for background estimation whenever possible, some background processes are difficult to estimate without MC simulation. These processes mimic the signal very closely, and therefore it is difficult to obtain a control region suitable for data-driven estimation. In such cases, the MC generators are important for the estimation of both shape and the normalization. The shape uncertainties are accounted for in the previous section by comparing several generators. For the normalization, we rely on published studies.

The Wt single top process is studied in [9] where scale uncertainty is estimated to be $\sim 3\%$. However, it is also shown that there is a larger uncertainty due to the p_{T} cut used for b-jet veto. As we do not use a b-jet veto in our analysis, we estimate the uncertainty from the variation due to change in b-jet veto over the range of cuts studied in the reference. The PDF uncertainty for Wt was calculated to be $\sim 2\%$. We estimate the overall uncertainty on Wt to be 8%. For diboson processes [10], the scale uncertainty is estimated to be $\sim 3\%$, the PDF uncertainty is $\sim 4\%$.

To calculate the uncertainties on our measurement due to theoretical uncertainty on the cross-sections, we vary all Monte-Carlo driven processes by 5% except Wt for which we vary its cross-section by 8%. As some parts of the systematics such as PDF uncertainty can be correlated, we conservatively vary the uncertainties for all processes in a fully correlated manner.

5.8   Drell-Yan Background ($\alpha_{Zee/Z\mu\mu}$)

Systematic uncertainty for the DY background (Section 3.1) is obtained by varying both the $E_{\mathrm{T}}^{\mathrm{miss}}$ boundaries and the Z window (separately and simultaneously). We move the lower $E_{\mathrm{T}}^{\mathrm{miss}}$ bound to 10 GeV and expand the Z window to cover 85–97 GeV. These two separate variations give two new estimates of the DY background, and making the shifts simultaneously gives a fourth estimate. We take the average of the largest and the smallest of the four numbers as our estimate, and half the difference between the two as the systematic uncertainty on this estimate. With this approach the systematic uncertainty covers all the values we obtain for different values of $E_{\mathrm{T}}^{\mathrm{miss}}$ / dilepton mass cut. We conclude that the systematic uncertainty is about 15% for both the ee and $\mu\mu$ channels for $200\,\mathrm{pb}^{-1}$. The systematics is sensitive to the number of events in different $E_{\mathrm{T}}^{\mathrm{miss}} - M_Z$ regions, and therefore we expect it to be higher for smaller datasets ($50\,\mathrm{pb}^{-1}$, $100\,\mathrm{pb}^{-1}$).

5.9   Jets Misidentified as Leptons (α_{fake})

The largest systematic associated with jets faking leptons arises from measuring the fake rate in control samples and extrapolating to the signal region. Fortunately, we have two different control samples, and we can look at the variation between the two control samples to estimate how much the fake rates varies in different types of events. The difference between the control samples may be smaller than the difference

between the control samples and the signal region. Out of caution we take twice the difference between the two predictions as our systematic.

Based on the current studies with QCD Monte-Carlo samples with limited statistics we estimate an uncertainty in the fake rates of 100% for the early data ($50 \, \text{pb}^{-1}$) for both electrons and muons; and of 50% for the muons and 100% for the electrons for $100 \, \text{pb}^{-1}$ and $200 \, \text{pb}^{-1}$.

5.10 Pile-up Effects

The exact beam conditions for first data are largely unknown, but if we assume that $200 \, \text{pb}^{-1}$ of data is produced then it is likely that the LHC will have reached luminosities on the order of $10^{32} \, \text{cm}^{-2}\text{s}^{-1}$. At this luminosity pile-up effects will become noticeable. To estimate the pile-up effects we assume luminosities of $10^{32} \, \text{cm}^{-2}\text{s}^{-1}$, and bunch spacings of 450ns, as these conditions produce the largest average number (~ 4) of proton-proton interactions per bunch crossing expected in early data. We compare fully simulated MC datasets that include hits from cavern background and multiple proton-proton interactions to the same datasets without pile-up effects. The effect of pile-up on the signal acceptance is estimated to be 8% for the ee channel, 5% for the $\mu\mu$ channel, and 2% on the $e\mu$ channel. However, there are several effects which compensate and we can't assess it with precision, so we don't include the pile-up systematics to our final estimates of the total systematic uncertainty.

6 Results

We present results based on the "Asimov dataset" [8], which represents the expectation of an ensemble of pseudo-data sets drawn from our nominal signal and background expectations. Here we have partitioning of the MC-based background and the "fake" background contributions as if we had a data-driven background estimate for background sources with fake leptons. All cross-section uncertainties are taken to fluctuate in a totally correlated manner; this does not include the signal nor the fake backgrounds, which will be extracted from data driven techniques. The measurements presented are based on the number of events with two or more jets in each channel. We consider an integrated luminosity of $200 \, \text{pb}^{-1}$ with a 20% uncertainty.

Table 6 shows the individual contributions to the relative uncertainty on the cross-section for each of the channels individually and in combination for $200 \, \text{pb}^{-1}$. The first row indicates the statistical uncertainty, the second row indicates the dominant uncertainty from the luminosity uncertainty, and the next several rows indicate uncertainties from individual sources of systematics (determined by fixing all other systematics to their nominal values). The final two lines indicate the effect from all systematics except the luminosity and all systematics together.

While the systematics are grouped such that the sources are uncorrelated, their impact on the cross-section measurement cannot be simply added in quadrature. For instance, if the luminosity is actually higher and the jet energy scale is lower than their nominal values, then the expected number of events may not be very different than the nominal prediction. The correlated effect on the measurement is summarized by a correlation matrix in the fitted parameters of the model (see Table 7).

The log-likelihood curves obtained from fitting each channel individually and combined are shown in Fig. 11. Note the asymmetric nature of the profile likelihood curve introduced by the systematics.

7 Conclusion

In this note, we studied the prospects for $t\bar{t}$ cross-section measurement using the dilepton final states. The emphasis was on the analysis of early data and therefore the object and event selection strategy was kept as simple as possible. The basic strategy was to use a well identified lepton pair and the remaining

Table 6: The individual contributions to the relative uncertainty on the cross-section expected for each of the channels individually and in combination for $200\,\mathrm{pb}^{-1}$. The uncertainties are listed as "negative uncertainty / positive uncertainty". The uncertainties are asymmetric and define a 68% confidence interval.

$\Delta\sigma/\sigma$ (%)	ee channel	$\mu\mu$ channel	$e\mu$ channel	combined
Stat only	-7.5 / 7.8	-6.0 / 6.2	-4.0 / 4.1	-3.1 / 3.1
Luminosity	-17.3 / 26.3	-17.4 / 26.2	-17.4 / 26.2	-17.4 / 26.2
Electron Efficiency	-4.5 / 5.0	0.0 / 0.0	-2.2 / 2.4	-1.9 / 1.9
Muon Efficiency	0.0 / 0.0	-4.6 / 5.2	-2.1 / 2.2	-2.2 / 2.3
Lepton Energy Scale	-0.3 / 1.6	-2.4 / 2.0	-0.5 / 0.5	-0.8 / 0.8
Jet Energy Scale	-3.4 / 3.2	-3.0 / 4.5	-2.5 / 2.5	-2.8 / 3.0
PDF	-2.1 / 2.3	-1.4 / 1.6	-1.6 / 1.8	-1.7 / 1.8
ISR FSR	-4.0 / 4.2	-3.6 / 3.7	-3.5 / 3.5	-3.6 / 3.7
Signal Generator	-4.7 / 5.4	-4.6 / 5.4	-4.7 / 5.3	-4.7 / 5.3
Cross-Sections	-0.3 / 0.3	-0.3 / 0.3	-0.3 / 0.3	-0.3 / 0.3
Drell Yan	-1.4 / 1.3	-2.2 / 2.2	-0.5 / 0.5	-0.8 / 0.9
Fake Rate	-9.7 / 9.5	-1.1 / 1.1	-6.2 / 6.2	-4.0 / 4.0
All syst but Luminosity	-12.7 / 13.9	-8.9 / 10.2	-9.4 / 10.2	-8.7 / 9.6
All systematics	-21.0 / 30.3	-19.3 / 28.3	-19.5 / 28.5	-19.3 / 28.1
Stat + Syst	-22.3 / 31.3	-20.2 / 29.0	-19.9 / 28.8	-19.5 / 28.3

Table 7: The correlation coefficients for the combined fit including all systematics. The diagonal elements are unity by definition. The non-zero off-diagonal elements indicate a correlated impact on the final observables.

	$\mathcal{L}$	$\frac{\sigma_{\mathrm{sig}}}{\sigma_{\mathrm{SM}}}$	α_{Zee}	$\alpha_{Z\mu\mu}$	$\alpha_{\epsilon(e)}$	α_{fake}	$\alpha_{\epsilon(\mu)}$	α_{PDF}	$\alpha_{cross\,sec.}$	α_{eES}	α_{IFSR}	α_{JES}	$\alpha_{\mu\mathrm{ES}}$	$\alpha_{\mathrm{MC\,model}}$
$\mathcal{L}$	1.00	-0.91												
$\frac{\sigma_{\mathrm{sig}}}{\sigma_{\mathrm{SM}}}$	-0.91	1.00	-0.02	-0.04	-0.06	-0.21	-0.16	-0.07	-0.01	-0.01	0.15	-0.14	-0.04	-0.22
α_{Zee}		-0.02	1.00		-0.02	-0.03	0.02					-0.01		
$\alpha_{Z\mu\mu}$		-0.04		1.00	0.04	0.12	-0.06			0.01		-0.04	-0.03	
$\alpha_{\epsilon(e)}$		-0.06	-0.02	0.04	1.00	-0.23	0.11	-0.02			0.01	0.02	0.04	
α_{fake}		-0.21	-0.03	0.12	-0.23	1.00	0.25	-0.03		-0.01	0.01	0.09	0.11	
$\alpha_{\epsilon(\mu)}$		-0.16	0.02	-0.06	0.11	0.25	1.00	0.02			-0.01	-0.04	-0.05	
α_{PDF}		-0.07			-0.02	-0.03	0.02	1.00						
$\alpha_{cross\,sec}$		-0.01							1.00					
α_{eES}		-0.01		0.01		-0.01				1.00		0.01	0.01	
α_{IFSR}		0.15			0.01	0.01	-0.01				1.00	0.01		
α_{JES}		-0.14	-0.01	-0.04	0.02	0.09	-0.04			0.01	0.01	1.00	-0.03	
$\alpha_{\mu\mathrm{ES}}$		-0.04		-0.03	0.04	0.11	-0.05			0.01		-0.03	1.00	
$\alpha_{\mathrm{MC\,model}}$		-0.22												1.00

(a) ee channel

(b) $\mu\mu$ channel

(c) $e\mu$ channel
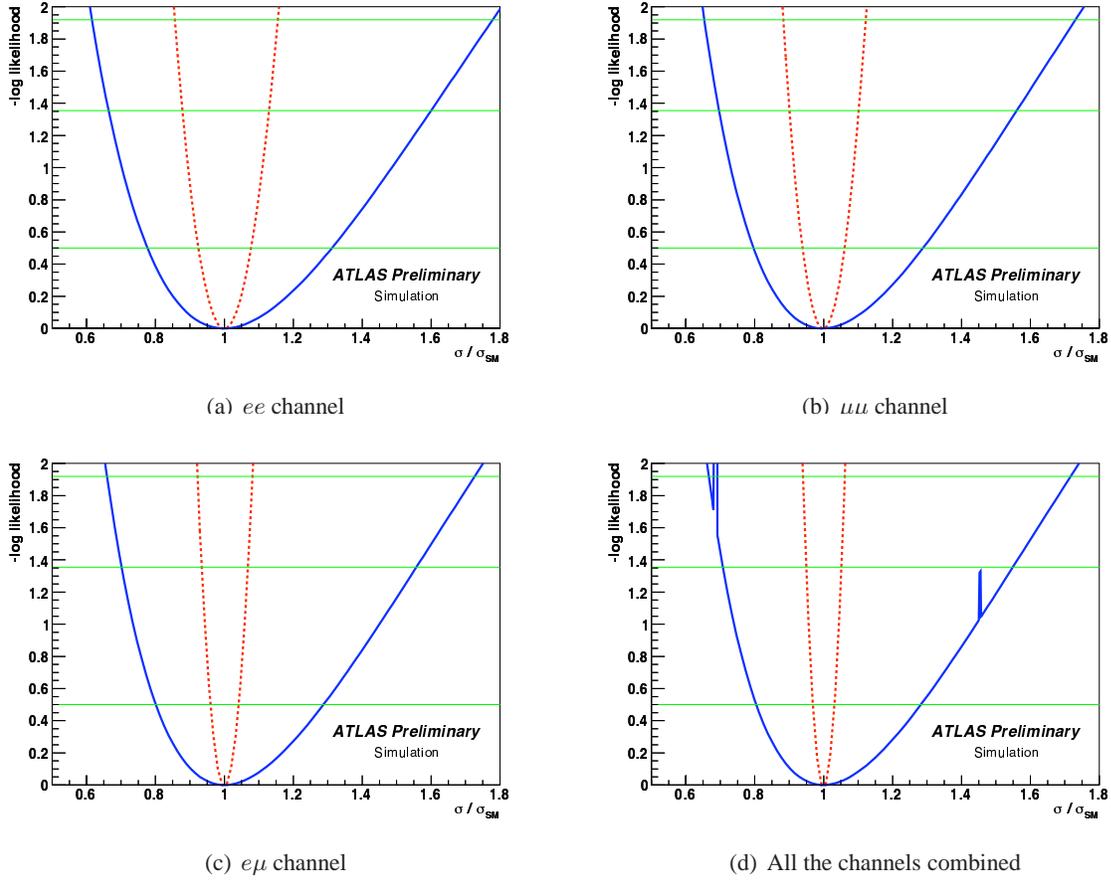
(d) All the channels combined

Figure 11: The log-likelihood curves for ee, $\mu\mu$ and $e\mu$ channels. The solid blue curve is the log of the profile likelihood ratio $-\log\lambda(\sigma_{sig})$, which includes all sources of systematics. The dotted red curve is the log of the likelihood ratio $-\log r(\sigma_{sig})$, which can be considered as including only statistical uncertainties. The horizontal green lines indicate 68%, 90%, and 95% thresholds (from bottom to top).

background was removed using the dilepton invariant mass, the E_T^{miss} and the jet multiplicity. The overall selection efficiency (S/B ratio) is $16.5\%(4.1)$, $26.1\%(3.8)$ and $26.5\%(5.5)$ for the ee, $\mu\mu$ and $e\mu$ channels respectively.

We studied data-driven methods for the estimation of background. In particular, strategies to determine Drell-Yan and fake background were developed. Uncertainties related to the methods were also estimated. The signal and Monte-Carlo based background were defined carefully to avoid any overlap with the data-driven components.

A range of potential uncertainties were studied in addition to the ones related to the data-driven methods. In particular, uncertainties on jet energy scale, lepton efficiency and Monte-Carlo model turned out to be the largest contribution to the systematics after the uncertainty on luminosity, which is by far the leading constraint on the measurement. On the other hand, with high selection efficiency and the large expected cross-section, statistical uncertainty will not dominate the final uncertainty once several tens of pb^{-1}s of data will be accumulated.

All uncertainties were combined by constructing a likelihood function for each channel. They were fit on the nominal prediction from Monte-Carlo samples and the final sensitivity was obtained from a profile likelihood ratio. The three channels were finally combined by performing a simultaneous fit incorporating the correlations between uncertainties.

In conclusion, the $t\bar{t}$ cross-section can be measured realistically using the first data expected from the LHC even at the 10 TeV collision energy. While the exact sensitivity depends on the systematic uncertainty extracted from the data, we can expect a competitive result from the first-year physics runs.

References

[1] C. Amsler et al. *Phys. Lett. B*, 667:1, 2008.

[2] P. Nason, S. Dawson, and R.K. Ellis. *Nucl. Phys.*, B303:607, 1988.

[3] W. Beenakker et al. QCD Corrections to Heavy Quark Production in $p\bar{p}$ Collisions. *Phys. Rev.*, D40:54, 1989.

[4] W. Bernreuther et al. *Nucl. Phys.*, B690:81 [arXiv:hep–ph/0403035], 2004.

[5] M. Cacciari, S. Frixione, L.M. Mangano, P. Nason, and G. Ridolfi. Updated predictions for the total production cross section of top and of heavier quark pairs at the Tevatron at the LHC. *JHEP*, 09:127 [arXiv:0804.2800], 2008.

[6] N. Kidonakis and R. Vogt. Next-to-next-to-leading order soft-gluon corrections in top quark hadroproduction. *Phys. Rev.*, D68:114014, 2003.

[7] S. Moch and P. Uwer. Theoretical status and prospects for top-quark pair production at hadron colliders. *Phys. Rev.*, D78:034003 [arXiv:0804.1476], 2008.

[8] ATLAS Collaboration. Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics. *arXiv:0901.0512*, CERN-OPEN(2008-020).

[9] J. Campbell and F. Tramontano. NLO corrections to Wt production and decay. *Nucl. Phys. B*, B276, hep-ph(0506289):109–130, 2005.

[10] M. Grazzini. Soft-gluon effects in WW production at hadron colliders [arXiv:hep-ph/0510337].