CMS Physics Analysis Summary

Contact: cms-pag-conveners-top@cern.ch 2009/07/08

Plans for an early measurement of the $t\bar{t}$ cross section in the electron+jets channel at $\sqrt{s}$ = 10 TeV

The CMS Collaboration

Abstract

We describe plans for a measurement of the $t\bar{t}$ cross section using semileptonic decays of top quark pairs in the electron channel that does not rely on b-tagging. Data-driven methods to estimate the contributions from QCD and W+jets background processes are presented. For QCD, we have explored a one-dimensional extrapolation method to estimate the QCD background using an electron isolation distribution. For W+jets, we have studied a template fit method using a discriminating variable. We demonstrate that given 20 pb^{-1} of LHC collision data at a center-of-mass energy of 10 TeV, it is possible to measure the $t\bar{t}$ cross section with a statistical uncertainty of 23%. The systematic uncertainty of this measurement is estimated to be 20%.

1 Introduction

In this note, we describe an analysis designed to measure the top quark pair production cross section in the electron+jets channel, $t\bar{t} \to be\nu\bar{b}q\bar{q}$, with CMS using early LHC data equivalent to 20 pb^{-1} at a center-of-mass energy of 10 TeV. With this integrated luminosity, we expect about 8300 $t\bar{t}$ pairs to be produced at the LHC [1]. We do not use b-tagging in this analysis.

The major backgrounds to this channel can be broadly divided into two kinds: (i) backgrounds with a real prompt electron, e.g. W+jets (where the W boson decays to electron and a neutrino), Z+jets (where the Z boson decays to an electron-positron pair) and single top production (where the top quark decays semileptonically to electron); (ii) backgrounds with fake or secondary electrons arising from QCD multijet events.

The two main backgrounds are expected to come from W+jets and QCD. A data-driven method is developed to estimate the size of the QCD background while a technique that combines Monte Carlo (MC) simulation and data-driven methods is developed to estimate the W+jets background.

The simulation of $t\bar{t}$ signal events was performed using a matrix element (ME) MC calculation, MADGRAPH [2] interfaced to PYTHIA [3] for the parton showering (PS) using the MLM matching prescription [4]. Background processes, W/Z+jets and single top, were also simulated using this combination of MADGRAPH+PYHTIA. Only leptonic decays of W/Z bosons were retained. The QCD background was simulated with PYTHIA where a high-statistics sample was produced which was filtered at the generator level for the presence of an electromagnetic object (e.g. an electron). We normalize all processes to their respective NLO cross sections [1, 5].

2 Event Selection

The experimental signature for the e+jets final state is an event which contain one high transverse momentum (p_T) electron, at least four jets and missing transverse energy ($\not\!\!E_{\mathrm{T}}$). We use a single electron trigger with a transverse energy (E_T) threshold of 15 GeV to record these events. We apply a "baseline selection" to all events passing this trigger. Subsequently, we explore two complementary sets of additional requirements in order to reduce events from background processes to acceptable levels.

2.1 Baseline Selection

The reconstruction of electrons [6] uses information from the pixel detector, the silicon strip tracker and the electromagnetic calorimeter (ECAL). Starting from "superclusters" (clusters of clusters) in the ECAL a supercluster-driven pixel-seed finding is then performed in order to initiate the trajectory-building in the inner tracker. A "Gaussian sum filter" [7] is then used to reconstruct the tracks of the electrons.

A set of shower shape variables is used for electron identification. These variables are intended to be used at startup since they are independent of the measured fraction of Bremsstrahlung and they are insensitive to tracker misalignment. The variables used are $\sigma_{\eta\eta}$, the shower shape in pseudorapidity (η), $\Delta\eta_{in}$ and $\Delta\Phi_{in}$, the difference in η and Φ between the track and the supercluster, and H/E, the ratio of energy deposited in the hadronic calorimeter (HCAL) to the energy deposited in the ECAL.

We require exactly one electron passing this electron ID to have $E_T > 30$ GeV, $|\eta| < 2.5$ (excluding electrons from the transition region between barrel and endcap detectors: $1.442 < |\eta| <$

1.560), and transverse impact parameter with respect to the beamspot (d_0) less than 200 μm. The cut on the pseudorapidity is motivated by the acceptance of the electron trigger.

We require these electrons to be isolated, making use of tracker and calorimeter isolation variables. Tracker isolation I_{trk} is calculated by summing the transverse momentum of tracks with at least 1 GeV in p_{T} within a cone of $\Delta R = 0.3$, where $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. ECAL isolation I_{ecal} is calculated by summing the transverse energy deposited in the crystals in a cone of $\Delta R = 0.4$. HCAL isolation I_{hcal} is calculated by summing the transverse energy of hadronic towers in a cone of $\Delta R = 0.4$. We combine these three variables into one combined relative isolation variable:

$$RelIso = \frac{I_{\text{trk}} + I_{\text{ecal}} + I_{\text{hcal}}}{E_{\text{T}}},$$

where E_{T} is the transverse energy of the electron. We require exactly one isolated electron with RelIso < 0.1. In CMS, electron identification efficiency and trigger efficiency will be measured using a Tag-and-Probe Method [8].

We reject events which contain one or more muons with $p_T > 20$ GeV/c, $|\eta| < 2.1$, RelIso$<$ 0.05, $d_0 < 200\ \mu$m, $\chi^2/d.o.f. < 10$ for the global muon track fit, and number of hits in the silicon tracker, $N_{\text{hits}} \geq 11$.

For the clustering of the jets the Seedless Infrared-Safe Cone (SISCone) algorithm [9] with a cone size of $R = 0.5$ is used. A detailed study on the performance of this algorithm for the CMS detector can be found in [10]. CMS plans to determine the Jet Energy Scale (JES) with a multi-step approach [11]. In this analysis the relative and the absolute corrections are applied. We require at least four jets in the range $|\eta| < 2.4$ with $p_T > 30$ GeV/c. Jets close to an identified electron ($\Delta R(j,e) < 0.3$) are considered to be electrons and removed from the jet list.

2.2 Additional Cuts to Reduce Z+jet Background

Since the electron identification criteria are relatively "tight", there are a significant number of di-electrons from Z decays which fulfill the requirement for exactly one electron because the second electron does not meet the "tight" cuts. In order to reduce this background, we have studied two approaches.

In the first approach (Option 1), we reject events with any additional electrons with $E_T >$ 20 GeV, RelIso$<$ 0.25, and within the range $|\eta| < 2.5$. We also do not apply any electron ID for these loose electrons. Applying this veto, we reject 56% of the Z+jets events with a 3% signal loss.

In the second approach (Option 2), we look for an additional reconstructed electron which has at least 20 GeV in E_{T}, $|\eta| < 2.5$ and passes loose electron identification. If the invariant mass of these two electrons is within 76-106 GeV/c^2, we reject the event.

2.3 Additional Cuts to Reduce QCD

After the soft electron veto (Option 1) has been applied, there is still a significant contribution from QCD multijet events (76 events compared with 208 signal events). Fake electrons from QCD come in four forms: fakes from charged pion/photon overlap, decays of hadrons containing b/c quarks to electrons, charge exchange background of charged pions to neutral pions, and fake electrons from converting photons. In order to further reduce these background processes, we again studied two approaches.

The first approach (Option 1) tightens the pseudorapidity requirement and only accepts events

with a tight electron within the barrel-region of $|\eta| < 1.442$. Since most of the material before the calorimeters in CMS is in the forward region, omitting the endcaps reduces the number of electrons from conversions considerably.

The second approach (Option 2) makes use of a dedicated photon conversion removal algorithm. When a photon converts, the conversion electrons will bend in the magnetic field, and since they have opposite charges they will bend in opposite directions in the ϕ plane. Our algorithm looks for pairs of such tracks. These tracks should come from a point from where the conversion took place. We require that the tracks have opposite charges. We calculate the closest 2D distance between the two tracks in the ϕ plane. If this distance is less than 0.04 cm and $|\Delta \cot(\theta)| < 0.03$, the electron is flagged as a conversion electron, and the event is rejected. The efficiency that can be achieved with the algorithm is around 50%, with a signal loss (determined from signal Monte Carlo simulations of isolated electrons) of 1.7%.

To suppress further the QCD (and Z+jets) background, the second approach also requires $\not\!\!E_T >$ 20 GeV, after correcting for the momenta of any reconstructed muons. In Figure 1 we show the distribution of $\not\!\!E_T$ for the various physics processes where the separation between events with real $\not\!\!E_T$ and events without real $\not\!\!E_T$ can be seen.

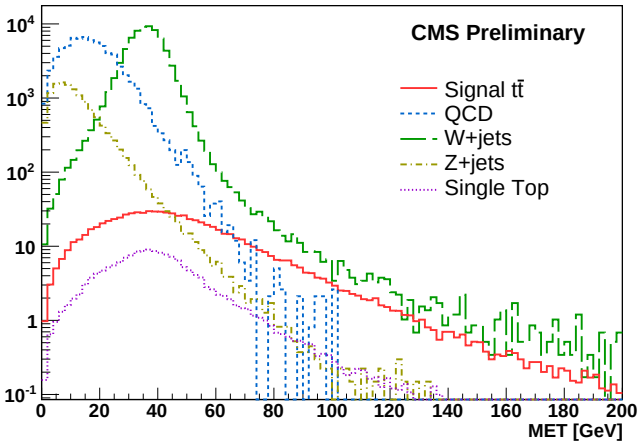

Figure 1: Missing transverse energy ($\not\!\!E_T$) distributions for events that pass the electron trigger and contain at least one electron satisfying the selection criteria described in the text (except isolation). Events normalized to 20 pb^{-1}.

With either approach, the QCD contamination to the e+jet sample is reduced to manageable levels. Experience with early data (e.g. how well $\not\!\!E_T$ is understood) will influence which approach is ultimately taken.

3 Expected Event Yields from Monte Carlo Simulation

In Table 1 we show the total number of selected events from each physics process after each cut (normalized to 20 pb^{-1}). The composition of the events as a function of jet multiplicity are shown in Figure 2. We illustrate how background and signal events are distributed for key kinematic variables in Figure 3.

	Cuts	$t\bar{t}$	W+jets	Z+jets	QCD	Single Top
	-	8280 ± 6	$9.1E5 \pm 265$	$8.4E+04 \pm 60$	$1.7E8 \pm 3.4E4$	1455 ± 2
	Trigger	4727 ± 5	$2.0E5 \pm 175$	$2.7E+04 \pm 42$	$3.4E7 \pm 1.9E4$	669 ± 2
	$\geq$1 Iso e	654 ± 2	$6.4E4 \pm 102$	$1.2E+04 \pm 29$	9030 ± 318	148 ± 1
	=1 Iso e	640 ± 2	$6.4E4 \pm 102$	8672 ± 25	9030 ± 318	146 ± 1
	Muon Veto	590 ± 2	$6.4E4 \pm 102$	8664 ± 25	9030 ± 318	143 ± 1
	$\geq$ 4 jet	215 ± 1	95 ± 4	46 ± 2	76 ± 20	10 ± 0
Option 1	Loose e Veto	208 ± 1	95 ± 3	20 ± 1	76 ± 13	10 ± 0
	$\lvert\eta\rvert < 1.442$	172 ± 1	57 ± 2	12 ± 1	31 ± 10	8 ± 0
Option 2	$\not\!\!E_T > 20$ GeV	188 ± 1	83 ± 4	34 ± 2	48 ± 15	9 ± 0
	Z Veto	186 ± 1	83 ± 4	29 ± 2	48 ± 15	9 ± 0
	Conv. Veto	183 ± 1	80 ± 4	28 ± 1	30 ± 14	9 ± 0

Table 1: Number of events passing each stage of the event selection normalized to 20 pb^{-1} as estimated from Monte Carlo Simulation. The uncertainties are the statistical errors from the corresponding Monte Carlo samples.

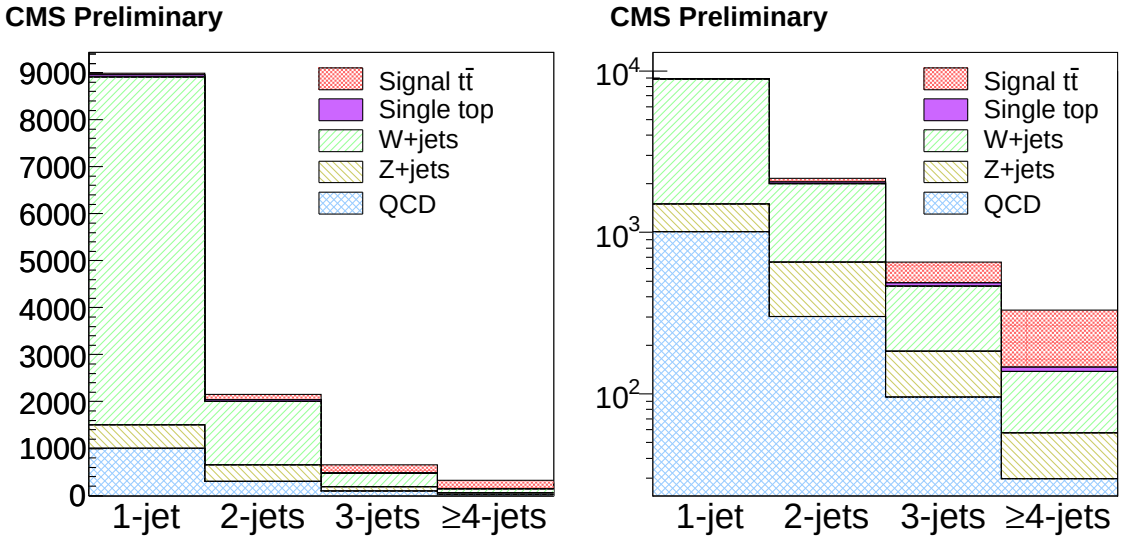

Figure 2: Expected number of signal and background events as a function of jet multiplicity normalized to 20 pb^{-1}: linear scale (left) and logarithmic scale (right). Events are after all selections (option 2) except jet multiplicity cut.

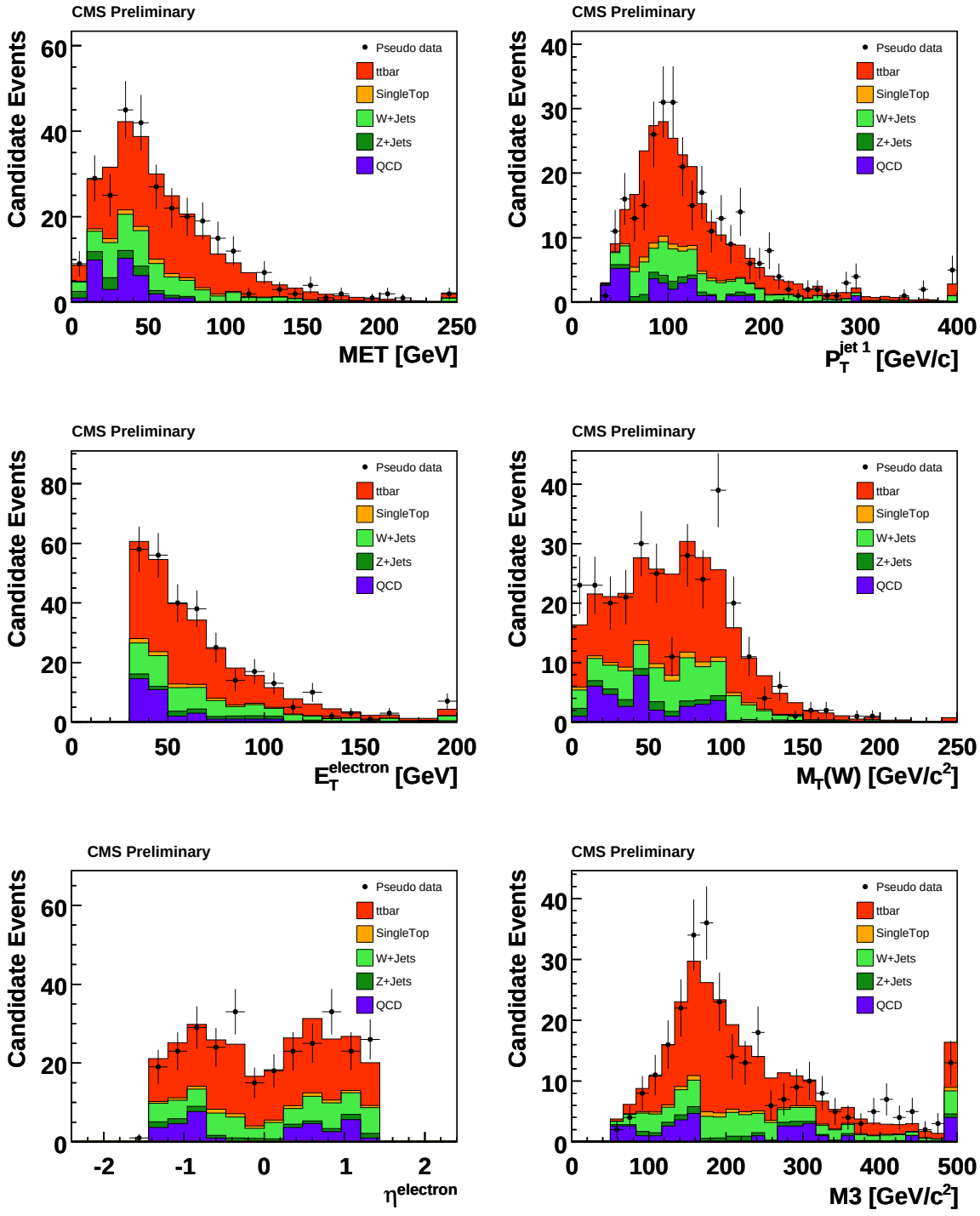

Figure 3: Top row: The left plot shows the $\not\!\!E_T$ distribution. On the right-hand side the p_T of the leading jet is shown. In the middle row the E_T of the electron is shown (left) together with the transverse W mass (right). In the bottom row, the plot on the left-hand side shows the η distribution of the electron, and the plot on the right-hand side shows M3, the invariant mass of the three jets with the largest vector sum of p_T. In all plots, the normalization of the different processes is according to the event yield in 20 pb^{-1} for Option 1. Pseudodata are an ensemble of randomly selected, Poisson fluctuated, simulated events drawn from each process.

4 Estimating the QCD background contribution from Data

We have developed a data-driven method to estimate QCD multijet backgrounds using the RelIso electron isolation variable. We exploit the fact that non-isolated electrons come mostly from QCD events, and hence we can use events with large isolation values (non-isolated) as a control region where QCD dominates. We fit the isolation distribution in the control region and extrapolate into the signal region characterized by small values of isolation, to obtain an estimate of the number of QCD events in the signal region.

While in this study we can test which function is best suited for this extrapolation by comparing various fits to a pure QCD MC sample where the behaviour in the signal region is known, this will not be possible with CMS data. Consequently, we have developed a method to determine the functional form from data. We find that by lowering the electron E_{T} threshold from 30 GeV to 20 GeV, inverting the $\not\!\!E_{\mathrm{T}}$ cut to $\not\!\!E_{\mathrm{T}}<20$ GeV, requiring that the scalar sum of transverse energies of electrons, jets and $\not\!\!E_{\mathrm{T}}$ be less than 100 GeV, and applying a tighter Z veto, a control sample with QCD purity $> 98\%$ can be obtained. We use this sample to find the best function with which to fit the signal sample. We tried various functions (Gaussian, Landau, polynomials of degree 3 and 4) to fit the selected events in this QCD dominated sample. The function that gives the most stable fit to the isolation distribution is the Landau function. In Figure 4 we show a fit of the Landau function to this QCD control sample in the range of RelIso from 0.2 to 1.0.

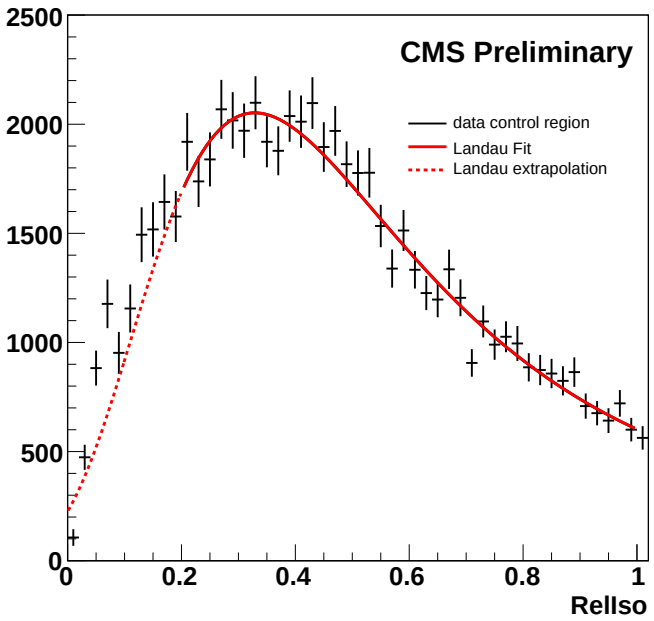

Figure 4: Fit of Landau function to RelIso distribution of events from a QCD dominated sample from 0.2 to 1.0. The dashed line below 0.2 is an extrapolation.

4.1 Results of QCD estimation

After applying the event selection described in Section 2 (except for the isolation and jet multiplicity cuts), we have only a limited number of QCD MC events remaining, especially in the higher jet-multiplicity bins. Since there may be insufficient QCD events to perform reliably the fit in the 3- and $\geq$4-jet bins, we perform the fit in the 1- and 2-jet bins and use information gained from these fits to improve the stability of the fit in the signal region. We found that if

we limit one of the parameters (MPV, or Mean Peak Value) to be within a range defined by the fitted values from the 1 and 2 jet bins, then we could obtain reliable estimates even in the 3 and 4 jet bins. This strategy can also be applied to data. The fit and extrapolation in the isolation distributions for events passing Option 2 of the event selection are shown in Figure 5 for the various jet multiplicity bins. A fit range from 0.2 to 1.0 is used. The results of the QCD estimation method are given in Table 2. We find the error on the fit underestimates the true uncertainty in the method and a $\pm 50\%$ uncertainty, derived by studying the robustness of the fit under varied conditions, is assigned instead.

We note that the RelIso extrapolation seems to systematically underestimate the true QCD contribution to the signal region. This underestimate arises from a difference in the RelIso shape in the high $\not\!\!E_{\mathrm{T}}$ signal sample when compared to the low $\not\!\!E_{\mathrm{T}}$ anti-selected control sample used to derive the Landau function described previously. This discrepancy is due to an excess of isolated electrons which are predominantly conversions.

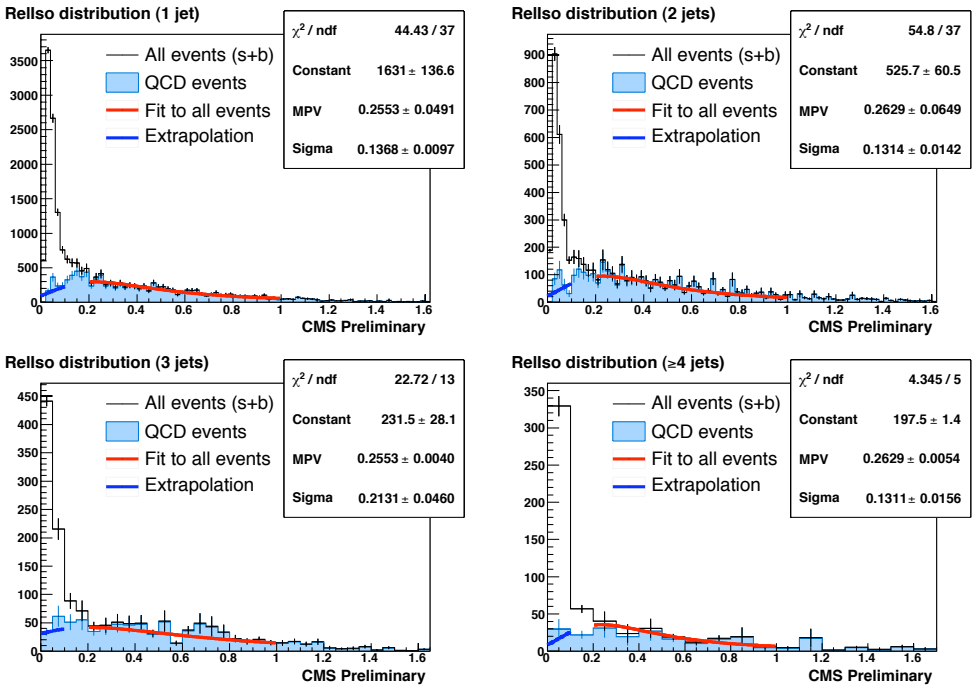

Figure 5: Distributions of the combined relative isolation of the electron in 1-, 2-, 3- and $\geq$4-jet bins. The lines show the Landau fit in the control region (0.2 to 1.0) and extrapolation into the signal region (0 to 0.1).

5 Separating Top Signal from W/Z+jets

Without b-tagging, we must rely on kinematic information to extract the top signal. To estimate the background contribution from W/Z+jets events, we use a template fit method which relies on a discriminant variable that has different shape in $t\bar{t}$ and W+jets events.

We use a variable called "M3" which is defined as the invariant mass of the three-jet combi-

	Signal region	
	True QCD $20\ \text{pb}^{-1}$	Estimate $20\ \text{pb}^{-1}$
1j	1007 ± 102	815
2j	301 ± 47	227
3j	96 ± 28	71
$\geq$4j	30 ± 14	17

Table 2: The QCD background estimation for 1-, 2-, 3- and $\geq$4-jet bins for events passing Option 2 of the event selection (RelIso requirement excluded).

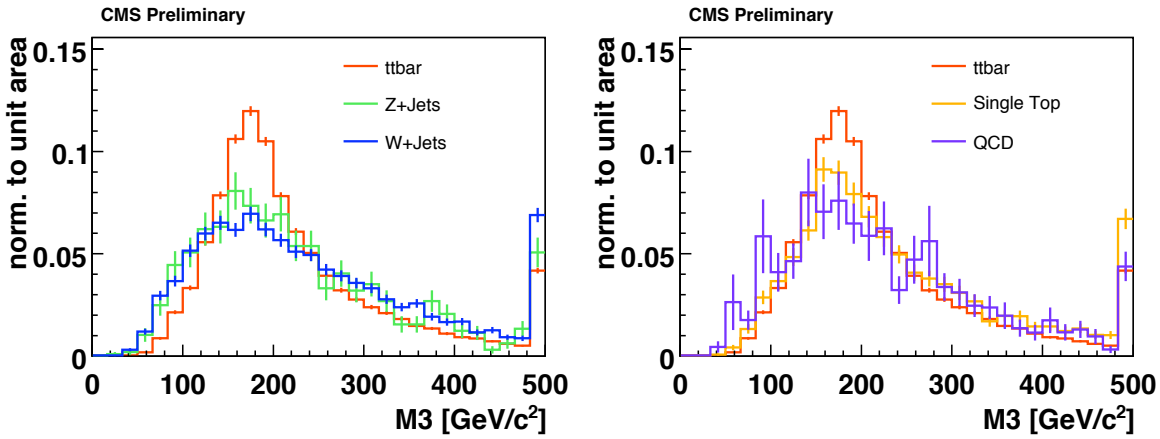

Figure 6: Comparison of M3 shapes for different processes passing Option 1 of the event selection. On the left hand side: the M3 distribution for $t\bar{t}$ as well as for W+jets and Z+jets. On the right hand side: the shape of the M3 distribution for $t\bar{t}$, single-top and QCD events.

nation which gives the highest vector-sum of transverse momenta of the jets. Only jets with $E_{\text{T}} > 30$ GeV with $|\eta| < 2.4$ are considered. A comparison of the M3 distribution for signal and background events passing Option 1 of the event selection is shown in Figure 6 where the discriminating power between signal and the W/Z+jets background can be seen.

We use an extended maximum likelihood method to perform the fit. We perform a four parameter binned likelihood fit to extract $N_{t\bar{t}}$, $N_{W/Z+jets}$, $N_{single\ top}$, and N_{QCD}. The M3 templates for the $t\bar{t}$, W/Z+jets, and single top (all processes) are taken from Monte Carlo simulation. Since the shape of Z+jets is similar to W+jet, we use the W+jets template to fit both W+jets and Z+jets events. A better approach would be to measure the shape of Z+jets in data and use that to fit both W/Z+jets. However with 20 pb^{-1}, there will not be enough Z+jets events. The QCD template is obtained from data in the non-isolated electron control region.

We impose a Gaussian constraint on the number of QCD and single top events. The QCD constraint is taken from the QCD estimation method described in the previous section. We set the mean of the Gaussian to the QCD estimate; the width is set to the 50% uncertainty in the QCD estimate. We do not use a data-driven method to estimate single top backgrounds. Therefore, we use the MC expectation as the mean of the Gaussian constraint. For the width of the Gaussian we use $\sqrt{\text{N}}$ where N is the expected number of single top events, but we also take into account the uncertainty in the theoretical cross section for single top, to which we attribute 30%. The overall factor used is $\sqrt{\text{N} + (0.3\text{N})^2}$.

We perform toy MC studies by generating pseudodata and fitting the templates to the pseudodata. In the generation of pseudodata, we first fluctuate the expected number of events for

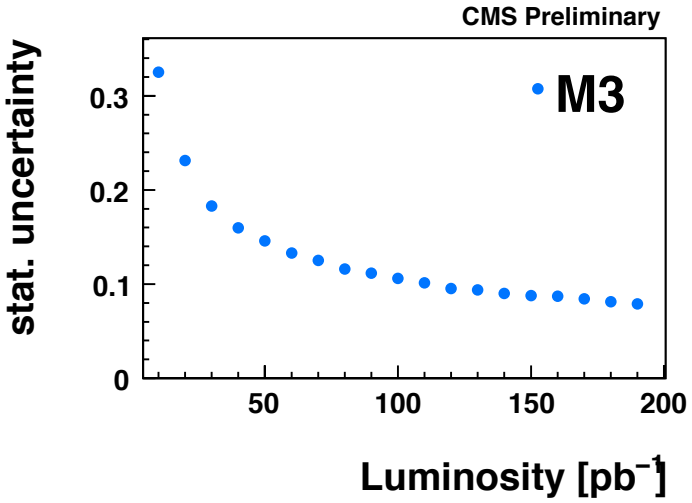

Figure 7: Expected statistical uncertainty of the $t\bar{t}$ cross section measurement as a function of integrated luminosity.

each physics process according to a Poisson distribution. We then generate pseudodata points for M3 by sampling randomly from the M3 distribution from MC. After that, we add up all the signal and background components to create a combined pseudodataset. Next we perform the fit on the combined pseudodataset. We do this to check the fit is free of bias and to ascertain the expected statistical uncertainty of the method, which for 20 pb^{-1} we estimate to be 23%. Figure 7 shows the improvement of the statistical error with increasing integrated luminosity; a statistical error of 10% can be reached within 90-100 pb^{-1}.

6 Sources and Estimates of Systematic Uncertainty

The impact of different sources of systematic uncertainties on the measured cross section is investigated with pseudodata ensembles. Systematic uncertainties can affect the number of selected events, i.e. the rate for a specific process, as well as the shape of the M3 distribution for this process. These effects can be correlated (or anti-correlated). We, therefore, evaluate rate and shape uncertainties simultaneously. The number of expected events for each physics process are fluctuated according to a Poisson distribution, but with a mean value which is derived from the systematically varied MC sample instead of from the nominal sample. For this number of expected events, we randomly sample the systematically modified M3 distribution.

6.1 Jet Energy Scale variation

In the absence of a jet p_T dependent model for the size of the JES uncertainty, a general uncertainty of 10% on the jet energy is assumed. The combined rate and shape uncertainty on the measured cross section due to the uncertainty in the JES is 15%.

6.2 Signal and Background Modelling

In order to estimate the systematic uncertainty due to the modelling of signal and background processes, different MC samples produced either with a different generator or with different settings for the modelling of $t\bar{t}$ and/or W+jets events were used. We used a PYTHIA $t\bar{t}$ sample

instead of the default $t\bar{t}$ MADGRAPH sample. The systematic error introduced on the extracted cross section is found to be 10%. We have also varied some of the parameters of the parton shower. These include: (1) the starting scale for the parton shower, which affects the hardness of additional jets; and (2) the scale used in the QCD running coupling of the shower, which affects the number of additional jets. The parameter variations chosen are conservative, and centered on values fit to other collider data. The resulting contribution to the estimated systematic error on the cross section is 3%.

In order to estimate the impact of the factorization scale in the production of W+jets events on the extraction of the $t\bar{t}$ cross section, the Q^2 scale is varied by a factor of 0.5 and 2.0 with respect to its default value. The systematic error on the $t\bar{t}$ cross section measurement due to the factorization scale uncertainty for W+jets is about 1%. This relatively small value results from an anti-correlation between the systematic uncertainty in the W+jets shape ($\sim$ 9%) with that induced in the acceptance.

In addition to the factorization scale uncertainty, the uncertainty on the ME-PS matching threshold for W+jets events was investigated. Therefore, the matching threshold for W+jets events was varied to 5 GeV/c and to 20 GeV/c compared to the default matching threshold of 10 GeV/c. The impact on the $t\bar{t}$ cross section measurement due to the uncertainty of the matching threshold for W+jets is in the order of 5%.

To assign a systematic uncertainty to the QCD shape we use MC (with contributions from all physics processes) in the control region to construct the QCD template, but the pseudodata is produced from only QCD events in the control region. The number of QCD events is constrained by the expected value. The shift induced in the fit corresponds to a $\pm$2% systematic uncerainty in the cross section.

To estimate the systematic uncertainty on the cross section measurement introduced by the uncertainty of the single top shape, an ensemble test is performed using a varied single top shape. The pseudodata is drawn from a shape, where the contribution of the tW-channel is enhanced by a factor of 2 while the contribution of the t-channel is kept constant, or where the contribution of the t-channel is enhanced by a factor of 2 and the tW-channel is kept constant. This leads to a varied relative fraction of the tW-channel to t-channel and thus to a varied single top shape. The systematic error on the cross section measurement introduced by this uncertainty is 1%.

Any uncertainty arising from the normalization of single top and QCD in the M3 fit is already accounted for in the statistical uncertainty quoted above. No such systematic uncertainty is assigned. We do, however, investigate the robustness of the fit with respect to the Gaussian constraints on the number of single top event and QCD event by varying the contributions for these processes by $\pm$50%. We find these variations to have a negligible effect.

6.3 PDF Uncertainty

The cross section uncertainty due to the imperfect knowledge of the proton parton density function (PDF) is evaluated using the 2*22 CTEQ6.6 [12] PDF sets and the LHAPDF [13] package using a reweighting procedure as described in [14]. The 2*22 PDFs are a set of uncorrelated positive and negative variations that span all sources of uncertainty in the CTEQ6.6 PDF. We arrive at a total uncertainty by summing the results of the individual variations in quadrature. We estimate an overall PDF uncertainty on the measurement of approximately 5%.

6.4 Summary of Systematic Uncertainties

A summary of systematic uncertainties on the measurement of the $t\bar{t}$ cross section can be found in Table 3. The systematic uncertainties quoted in the Tables are the largest deviations in order to have a conservative and symmetric estimation of the systematic uncertainty. The total uncertainty is calculated as the square root of the quadratic sum of the single uncertainties. One more systematic uncertainty is introduced by the uncertainty on measurement of the integrated luminosity, assumed to be 10%, is considered separately.

	Relative Systematic Uncertainty
Jet Energy Scale	15%
$t\bar{t}$ MC Generator	10%
$t\bar{t}$ ISR/FSR uncertainty	3%
W+jets MC Factorization Scale	1%
W+jets MC Matching threshold	5%
Shape uncertainty of Single Top	1%
Shape uncertainty of QCD	2%
PDF uncertainty	5%
Total	20%

Table 3: Systematic uncertainties on the estimation of the cross section. For each uncertainty we symmetrized the effect on the cross section by quoting the larger deviation. The total uncertainty is calculated as the square root of the quadratic sum of the single uncertainties.

7 Measuring the Top Cross Section

After the fit described above, we can make a measurement of the $t\bar{t}$ production cross section. To calculate the cross section we use the following expression:

$$\sigma_{t\bar{t}} = \frac{N_{t\bar{t}}^{\text{fit}}}{A \cdot \varepsilon_{t\bar{t}} \times \int \mathcal{L} dt}$$

where $N_{t\bar{t}}^{\text{fit}}$ is the number of observed signal events obtained from the M3 fit, A is the signal acceptance and $\varepsilon_{t\bar{t}}$ is the signal efficiency obtained from Monte Carlo simulation, and $\int \mathcal{L} dt$ is the integrated luminosity.

8 Improving the Signal-to-Background Ratio with b-tagging

If b-jets can be reliably distinguished from light quark jets (b-tagging), we can significantly increase S/B over that presented in the current analysis since top pairs produce two b-jets in the final state whereas the dominant backgrounds described in this note (QCD, W+jets) do not. We have not relied on such techniques in this analysis, to be prepared for the possibility that b-tagging will not be available in the first 20 pb^{-1} of data taking.

9 Conclusion

We have described an analysis to measure the cross section of $t\bar{t}$ production in the semileptonic decay of the top quarks in the electron channel assuming 20 pb^{-1} of data at $\sqrt{s}$ of 10 TeV. Data-

driven methods to estimate QCD and W/Z+jets backgrounds have been developed. The main irreducible background is expected to come from W+jets events.

Our studies show that it is possible to measure the cross section with a statistical uncertainty of 23% for 20 pb^{-1} of data. We have assessed a number of important sources of systematic uncertainties and estimated a total systematic uncertainty on the cross section measurement to be 20%.

References

[1] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason, and G. Ridolfi, "Updated predictions for the total production cross sections of top and of heavier quark pairs at the Tevatron and at the LHC," *JHEP* **09** (2008) 127, arXiv:0804.2800. doi:10.1088/1126-6708/2008/09/127.

[2] J. Alwall et al., "MadGraph/MadEvent v4: The New Web Generation," *JHEP* **09** (2007) 028, arXiv:0706.2334.

[3] T. Sjostrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual," *JHEP* **05** (2006) 026, arXiv:hep-ph/0603175.

[4] S. Hoche et al., "Matching parton showers and matrix elements," arXiv:hep-ph/0602031.

[5] S. Frixione and M. L. Mangano, "How accurately can we measure the W cross section?," *JHEP* **05** (2004) 056, arXiv:hep-ph/0405130.

[6] S. Baffioni et al., "Electron reconstruction in CMS," *Eur. Phys. J. C* **49** (2007) 1099–1116.

[7] R. Fruehwirth , *Comput. Phys. Commun.* **100** (1997) 1.

[8] **CMS** Collaboration, "Measuring Electron Efficiencies with Early Data," *CMS Physics Analysis Summary* **EGM-07-001** (2007).

[9] G. P. Salam and G. Soyez, "A Practical Seedless Infrared-Safe Cone jet algorithm," *JHEP* **05** (2007) 086.

[10] **CMS** Collaboration, "Performance of Jet Algorithms in CMS," *CMS Physis Analysis Summary* **JME-07-003** (2008).

[11] **CMS** Collaboration, "Plans for Jet Energy Corrections at CMS," *CMS Physics Analysis Summary* **JME-07-002** (2008).

[12] P. M. Nadolsky et al., "Implications of CTEQ global analysis for collider observables," *Phys. Rev.* **D78** (2008) 013004, arXiv:0802.0007. doi:10.1103/PhysRevD.78.013004.

[13] M. R. Whalley, D. Bourilkov, and R. C. Group, "The Les Houches Accord PDFs (LHAPDF) and Lhaglue," arXiv:hep-ph/0508110.

[14] **CMS** Collaboration, "Physics Technical Design Report, Vol. 2: Physics Performance," *J. Phys G: Nucl. Part. Phys.* **34** (2007) 995–1579.